Analyse spatiale et multi-échelle de la distribution des bactéries dans le sol et les sédiments

Jean-Sebastien Beaulne

ÉCOLE CENTRALE LYON

THÈSE

Présentée devant :

ÉCOLE CENTRALE DE LYON

Pour obtenir le grade de :

DOCTEUR

De l'école doctorale :

Électronique, électrotechnique, automatique

UNIVERSITÉ DE LYON

Spécialité : Microbiologie Environnementale

Soutenue publiquement le 20/11/2015

Par

Jean-Sébastien Beaulne

Multi-scale spatial analysis of microbial distribution in soil and sediment

Analyse spatiale et multi-échelle de la distribution des bactéries dans le sol et les sédiments.

Jury :

Timothy M. Vogel Professeur, Université Claude Bernard Lyon 1, Ecole Centrale de Lyon, Ecully, France	Directeur de Thèse
Joshua P. Schimel Professor, University of California, Santa Barbara, USA	Rapporteur
Soren J. Sorensen Professor, University of Copenhagen, Copenhagen, Denmark	Rapporteur
Alban Ramette Doctor, institute of social and preventive medicine, Berne, Suisse	Examinateur
Pascal Simonet DR CNRS, Ecole Centrale de Lyon, Ecully, France	Examinateur

Remerciements

La vie propose parfois de belles surprises. Il y a un peu plus de 3 ans je n'avais pas en tête de faire une thèse en Microbiologie en France. La première personne que je tiens à remercier, c'est celle qui m'a offert cette possibilité. Merci Tim, quand j'ai reçu ton courriel il y a 4 ans, je n'ai pas hésité à plonger dans l'aventure, laissant derrière moi travail et pays (Québec). Avec le recul je dois dire que c'est l'une des meilleures décisions que j'ai prise. Les rencontres que j'ai pu faire au cours de ces années, les défis scientifiques et humains auxquels j'ai dû faire face, les connaissances que j'ai acquises, les pays que j'ai découvert et toute les expériences extraordinaires que j'ai pu vivre découlent de ce petit courriel reçu un soir de janvier 2012. Alors, merci Tim ! Encore plus qu'un directeur de thèse, j'ai aussi rencontré une personne humainement et scientifiquement exceptionnelle ! Bien évidemment j'aimerais remercier mes collègues et amis que j'ai eu la chance de cotoyer durant la thèse. Merci Catherine, Sebastien, Laura, Lorrie, Joseph, Jerem, Christophe, Alban, Alexianne, David, Agathe. Les discussions que nous avons pu avoir, qu'elles soient philosophiques, scientifiques, abstraites ou ludique, ont été extrèmement enrichissantes et divertissantes. Merci aussi à Richard, discuter avec toi c'est toujours un grand plaisir, que l'ont parle de pêche ou de pinard. Merci à Pascal, je me souviendrai toujours de cette Hermitage 1986. Cotoyer un grand homme de la science comme toi fut un privilège. Merci également à tout les employé(e)s et stagiaires de l'Ecole Centrale avec qui j'ai eu la chance de passer un peu de temps. Merci à l'Union Européenne pour le financement de ma thèse (projet Marie Curie, Trainbiodiverse). Merci aussi a David Bowie qui m'a accompagné dans la rédaction.

D’un point de vue plus personnel, j’aimerais remercier Gaïdig, sans qui la vie perd de sa saveur. Merci d’être à mes cotés, traverser la vie avec toi c’est le plus beau des voyages ! Merci à mes amis en France sans qui je n’aurais pas pu décompresser et traverser les moments plus difficiles. Merci Fred et Steph, Aurélie et Antoine, Matthieu et Julia, Kevin et Océanne. Merci à mes amis au Québec, mes voyages de pêches annuelles avec vous m’ont permis de me ressourcer quand j’en avais besoin. Merci Julien, Fred, Gab, Oli, Seb et Steph. Merci aussi à Marc, je n’oublierai jamais que tu as été le premier à me donner ma chance. Merci à Matthieu et Serge, combien de fois j’ai rêvé d’apparaître au Benelux avec vous. Merci à mes frères et ma sœur, vous avez toujours été une source de motivation et de courage pour moi. Finalement, merci à mes parents, merci d’avoir cru en moi et de m’avoir tout donné pour que je réalise mes rêves. Ce doctorat il est pour vous.

Résumé

Les bactéries ont colonisé toutes les niches écologiques de la planète. Plus précisément, les sols sont l'hôte de la plus grande biodiversité terrestre, la faune microbienne. Cette grande diversité de bactéries et leur relative ubiquité rendent difficile l'identification des variables contrôlant la distribution spatiale des bactéries vivant dans le sol. Comme les bactéries du sol jouent un rôle important dans les grands cycles biogéochimiques globaux, il est important de mieux comprendre les variables qui peuvent influencer la composition bactérienne des sols. Dans cette thèse, nous émettons l'hypothèse que l'hétérogénéité de la composition de la communauté bactérienne apparaît à la même échelle spatiale que l'hétérogénéité des propriétés physico-chimiques du sol. Afin de comprendre la relation entre la composition bactérienne des sols (à l'échelle d'une carotte de sol jusqu'à l'échelle d'une région entière du nord de la France) et les paramètres physico-chimiques du sol à différentes échelles spatiales, nous allons utiliser une approche intégrant des données issues d'analyses SIG (Système d'Information Géographique), d'analyses physico-chimiques du sol et d'analyses des communautés bactériennes du sol. A travers une suite de trois expérimentations, nous allons répondre à trois questions: Es-ce qu'une pression environnementale uniforme à une plus grande échelle (cm) peut atténuer l'hétérogénéité microbienne à micro-échelle? Es-ce que les variables ayant une distribution spatiale suivant un gradient géographique sont des variables structurant fortement la distribution spatiale des bactéries à l'échelle de ce même gradient? Est-ce que certains bio-indicateurs à grandes échelles peuvent intégrer des groupes de variables pour modéliser la distribution des bactéries pour une région entière ?

Mots clés : Distribution spatiale, Bactéries, Sols, Métagénomique, SIG

Abstract

The bacteria have colonized all the niches of the planet. Specifically, soils are home of the largest terrestrial biodiversity, microbial fauna. This great diversity of bacteria and their relative ubiquity make it difficult to idendified variables driving the spatial distribution of bacteria living in the soil. As soil bacteria play a significant role in the main global biogeochemical cycles, it is important to better understand the variables that can influence bacterial composition of soils. In this thesis, we hypothesize that heterogeneity of the bacterial community composition appears at the same scale level as the heterogeneity of soil physicochemical properties. In order to understand the relationship of bacterial composition of soils (from core experiment to field study in large region in the northern France) and soil factors at different spatial scales, we will use an approach coupling GIS tools, soil physico-chemical analysis and 16S rRNA gene NGS. With Three set of experiment we will answer three questions: Can a uniform environmental pressure at a larger scale (cm) overcome microbial micro-scale heterogeneity? Are geographical gradients strong drivers of the microbial community structure at the scale of the gradient? Do large-scale geographical features that integrate groups of parameters model the differences in microbial community structure for an entire region?

Key words: Spatial distribution, Bacteria, Metagenomics, Soils, GIS

List of scientific publication

Oral Presentation

- 2nd Thünen Symposium on Soil Metagenomics, 11-13 December 2013, Braunschweig, Germany
- 6th Annual Argonne Soil Metagenomics Meeting, October 1-3, 2014, St-Charles, Illinois, USA
- Indo-French Workshop on Environmental Biotechnology, 9-10 Febuary, 2015, Bhubaneswar, Odisha, India
- Ecology of Soils Microorganisms, November 29 – December 3, 2015, Prague, Czech Republic (abstract accepted for oral presentation)
- Invited Lecturer, Biotechnology Research Institute of Canada, 29 September 2012

Poster Presentation

- The 14th International Symposium on Microbial Ecology, ISME14, 19-24 August, 2012, Copenhagen, Denmark
- 12th Symposium on Bacterial Genetics and Ecology (BAGECO), 9-13 June, 2013, Ljulbjana, Slovenia
- 3th Symposium on Bacterial Genetics and Ecology (BAGECO), 14-18 June, 2015, Milan, Italia

Scientific Paper (submitted)

- Geographic Information System (GIS) Analyses of Large-Scale Spatial Soil Bacterial Diversity, Submitted in October 2015, SBB
- Spatial and Temporal Analysis of Bacterial Diversity in Lake Sediment, submitted in October 2015, ISME

List of captions

Chapter 2

Chapter 3

Chapter 4

Table of Content

Synthèse en français

Introduction

La répartition spatiale des différentes espèces microbiennes et leurs fonctions dans les sols sont encore mal comprises, malgré un travail considérable sur la visualisation et le nombre de différents échantillons de sols analysés. Une partie de la difficulté est due aux limitations techniques et aux coûts liés à l'exploration des micro-organismes à différentes échelles spatiales. D'autre part, les macroorganismes ont été bien étudiés afin de savoir comment ils se distribuent dans l'espace et le temps, et quelles sont les variables clefs contrôlant cette distribution temporelle et spatiale. Seulement quelques principes écologiques tirés des recherches effectuées sur les macroorganismes, peuvent être transposés aux micro-organismes dûs à la définition d'espèces différentes et de capacités singulières comme le transfert horizontal de gènes. Les bactéries ont existé et évolué sur terre depuis des milliards d'années. Pendant cette période, et considérant que leur taux d'évolution est plus élevé que les macroorganismes en raison de leur taux de reproduction élevé (par multiplication ou par scissiparité) et à d'autres stratégies d'adaptation comme le transfert horizontal de gènes, ils ont évolué et se sont diversifié jusqu'à atteindre 10 millions d'espèces (Sogin et al., 2006). Pour accéder à ce grand réservoir génétique hautement diversifié, nous sommes confrontés au problème du grand nombre de microorganismes en faible abondance, mais presque toujours présent («low abundance biosphere"). Dans les sols, la faible abondance relative est amplifiée par le nombre d' «espèces» individuelles trouvées dans un gramme de sol. Certaines estimations suggèrent que le nombre d' « espèces » est aussi élevé que 10^7 par gramme de sol (Gans et al., 2005). Avec la longue histoire des bactéries sur terre, elles ont eu tout le temps nécessaire pour coloniser toutes les niches dans les sols

de la planète entière. Le transport éolien, le transport par l'eau et la motilité n'ont eu, depuis 3,5 milliards d'années, qu'à déplacer les bactéries de 1 cm par an en moyenne pour que celles-ci atteignent une distance cumulée de 35 000 km, ce qui est près de la circonférence de la terre (40 075 km). Sans considérer le transport éolien et la dispersion dans l'eau, même les frontières continentales n'ont pu limiter la dispersion spatiale car les bactéries étaient déjà là bien avant le super continent Pangée, rompu il y a 290 millions d'années. Donc, les questions concernant « is everything is everywhere » ont des dimensions temporelles et spatiales importantes. Déjà, à micro-échelle, il y a des quantités considérables d'espace vide de vie donc, clairement à cette échelle « everything is not everywhere ». Mais, à des échelles plus grandes, on trouve suffisamment de bactéries et de fonctions associées pour que ces dernières aient une influence importante sur les cycles biogéochimiques globaux (azote, carbone, pollution), contribuent ou dégradent la santé des plantes, et même aient un impact sur les changements climatiques. Il est donc important de mieux quantifier le rôle des bactéries dans les différents cycles globaux et, pour ce faire,connaître les variables qui influencent la distribution spatiale des bactéries et leurs activités à différentes échelles spatiales et temporelles.

Dans le premier chapitre de cette thèse, nous allons discuter des différentes variables qui ont été proposées comme contrôlant la distribution spatiale des communautés bactériennes du sol. Plusieurs variables ont été identifiées mais, leurs influences sur la distribution spatiale des bactéries opèrent à différentes échelles spatiales. À l'échelle bactérienne, la micro-échelle, la notion de micro-habitat (les réseaux d'air et d'eau, la granulométrie des agrégats, les pores) et des variables chimiques contrôlent la présence des bactéries (Grundmann et al. 2004, Franklin &

Mills 2003, Young et al. 2008). Ces caractéristiques physiques et chimiques sont difficiles à extrapoler à des échelles plus grandes; échelles qui sont plus pertinentes pour les études environnementales (de l'échelle du champ à l'échelle globale). En analysant les échantillons (de plusieurs mg) de sol nous détruisons les références micro-spatiales s'y rattachant, un peu comme si on mixait une forêt entière et que l'on essayait ensuite de comprendre la distribution spatiale d'origine des végétaux à partir des données génétiques et génomiques tirées de l'homogénat. A l'échelle méso (l'échelle du champ, de quelques mètres à quelques kilomètres) des études ont identifié la distribution des variables physico-chimiques, la couverture végétale et la disponibilité des nutriments comme étant des paramètres important de la distribution spatiale des micro-organismes, souvent en associant structure / fonction (Fierer et al. 2009, Berg & Smalla 2009). À une échelle encore plus grande (mondiale, continentale régionale), les gradients de températures, gradients d'élévations, les variables météorologiques, le substrat géologique, la couverture du sol, le pH, et les perturbations anthropiques ont été proposés comme variables contrôlant la structure de la communauté microbienne (Lauber et al. 2009, Fierer et al. 2012, Griffiths et al. 2011, Ge et al., 2008). Dans la deuxième partie du chapitre 1, nous allons essayer de répondre à la question "quelle est la taille d'un métagénome du sol". Récemment, le développement rapide de méthodes de biologie moléculaire a ouvert l'accès à plus de profondeur de séquençage, et donc, possiblement, une meilleure compréhension de la structure spatiale des communautés microbiennes et de leurs fonctions, notamment grâce à des approches de métagénomiques (Delmont et al. 2012).

Dans le chapitre deux, nous avons essayé de répondre à la question "Es-ce qu'une variable (contamination avec diesel) ayant une distribution à plus grande échelle (cm)

peut en partie masquer l'hétérogénéité à micro-échelle?" Pour ce faire, nous avons induit un changement sur la moitié d'une carotte de sol en ajoutant 200 ml de diesel. Après 14 jours d'incubation des deux demi carottes de sol (avec et sans diesel) nous avons analysé la communauté bactérienne pour voir si les changements induits à l'échelle centimétrique (même échelle que le sous-échantillonnage) effaçaient en partie l'hétérogénéité à micro échelle. Notre hypothèse étant que, si l'ampleur de la perturbation correspond à l'échelle de l'échantillonnage, nous devrions observer une transition vers des micro-organismes capables d'utiliser le substrat "diesel" et ainsi surpasser la micro-hétérogénéité et devenir la variable structurante de la distribution spatiale. Ces questions sont importantes pour l'analyse spatiale à plus grande échelle. Mesurer une transition rapide dans la communauté microbienne (14 jours) pourrait permettre de faire abstraction de la succession des couvertures du sol lors d'analyses de distribution à plus grande échelle. Il est également essentiel d'avoir des stratégies d'échantillonnages conformes aux échelles des variables qui influent sur la structure de la communauté.

Dans le chapitre 3, nous essaierons de répondre à la question «Est-ce que certains bio-indicateurs à grandes échelles peuvent intégrer des groupes de variables pour modéliser la distribution des bactéries pour une région entière". Nous avons choisi une région du nord de la France (250 km par 250 km). Un vaste ensemble de paramètres physico-chimiques ont été mesurés pour chaque point de prélèvement (64). A l'aide d'outils d'analyses spatiales et de logiciels SIG (Système d'Information Géographique), nous avons comparé l'approche statistique sur des données physico-chimiques et l'analyse spatiale de couverture du sol, pour analyser la structure de la communauté bactérienne. Notre hypothèse étant que les données physico-chimiques peuvent être corrélées avec la structure de la communauté, mais

que les systèmes d'informations géographiques fournissent des données essentielles sur les processus environnementaux dynamiques tels que le ruissellement et la couverture du sol, qui peuvent influer sur la structure des communautés en intégrant un ensemble de paramètres physico- chimiques.

Enfin, dans le chapitre 4, nous avons testé l'hypothèse que «les variables avec des gradients géographiques à l'échelle de la taille de l'échantillonnage sont les variables influençant la structure de la communauté à cette même échelle". Nous avons décidé de faire l'étude sur les sédiments de lac Chilika (Inde), le deuxième plus grand lac d'eaux saumâtres au monde. Le lac a un gradient de salinité allant de l'eau douce à l'eau marine. L'utilisation d'un ensemble de données physico-chimiques, l'analyse des gènes ARNr 16S et des analyses spatiales avec logiciel SIG,nous a permis d'identifier des variables clefs de la distribution spatiale des bactéries dans ce système particulier (par exemple : la salinité, les flux hydrauliques). En plus, nous avons pu décrire pour la première fois, en considérant les dimensions spatiales et temporelles, la distribution des populations microbiennes habitant ou transitant dans les sédiments de ce lagon d'eaux saumâtres.

Chapitre 1

La vie est présente sur terre depuis 3,5 milliards d'années. La longue présence et l'évolution des bactéries leurs ont permis de coloniser tous les niches de la planète. La très grande diversité de bactéries, comparée à celle des macro-organismes, est probablement due à cette longue période de temps et au fait que leur reproduction est beaucoup plus rapide. Depuis le début de l'écologie microbienne, nous abordons la question "es-ce que tout est partout?". Basse-Becking, en 1934, était le premier à tenter une réponse «tout est partout, mais l'environnement sélectionne". Il a été

montré, grâce à l'émergence des analyses métagénomiques, que les sols sont toujours très diversifiés mais que rien n'a une empreinte fonctionnelle plus similaire à un sol qu'un autre sol (Delmont et al., 2011). Pour comprendre comment les micro-organismes des sols sont organisés spatialement et qui est où et quand, nous devons déterminer les variables qui influencent la distribution spatiale des bactéries. Comprendre le lien entre les paramètres physico-chimiques et la structure de la communauté des sols, nous permettrait de modéliser la distribution spatiale globale des bactéries dans les sols et ainsi, de pouvoir modéliser ce qui se passe dans un environnement changeant.

Dans cette revue, nous allons décrire les différentes variables de la distribution spatiale des bactéries dans les sols ayant été identifiées dans la littérature. Nous allons déterminer à quelles échelles se produisent ou sont visibles ces variables. Différentes variables doivent être considérées pour une étude de terrain, en fonction de l'échelle à laquelle nous travaillons, et en fonction du maillage de l'échantillonnage. Tout se produit à la micro-échelle, l'échelle des bactéries, mais il est presque impossible de modéliser pour de grandes surfaces à partir de données extrapolées de la micro-échelle, en raison de la très forte hétérogénéité dans les paramètres environnementaux des sols. Certaines variables ont été identifiées comme des indicateurs de structures des communautés microbiennes, de l'échelle centimétrique à l'échelle continentale.

Indicateur à grande échelle

Gradient de température

Le premier paramètre à l'échelle globale sur lequel nous devons nous concentrer est le gradient de température. La plupart des macroorganismes ont des limites de dispersions partiellement en lien avec le gradient de température. Dans le monde macro, seulement l'espèce humaine a pu coloniser chaque environnement de ce gradient de température. Mais, en fait, cette possibilité est principalement due à notre capacité à modifier notre environnement (air conditionné, chauffage ...) et à se protéger artificiellement avec des vêtements et autres équipements thermiques.

Dans le monde micro, le lien entre la température et la distribution spatiale des micro-organismes n'est pas aussi évident. Il a été démontré que, contrairement au monde macro, les micro-organismes sont très diversifiés, même dans les régions nordiques (Neufeld et al., 2005) et que la structure des communautés microbienne qu'on y observe n'est pas si différente de ce que nous pouvons observer dans les sols de régions plus chaudes (Zhang et al., 2005 , Chu et al., 2010). Récemment, même dans la neige de l'arctique, une diversité importante de bactéries a été identifiée (Larose et al. 2010, Macario et al. 2014). A l'opposé, les micro-organismes sont également présents et diversifiés dans les environnements les plus chauds sur terre, là où aucun macroorganisme ne peut survivre, comme à proximité de la dorsale océanique (Huber et al., 2003).

Dans le contexte des changements climatiques, certaines études ont mis l'accent sur les conséquences d'une augmentation de la température (2 à 5 ° C) dans les sols etl'effet sur la structure de la communauté. Un changement vers une abondance relative plus élevée des champignons comparée à celles des bactéries a été observé après 2 à 5 ans d'augmentation artificielle de la température (Castro et al., 2010, Zhang et al., 2005). Une étude plus longue (12 ans d'augmentation de 5 ° au-dessus

T ° ambiante), a montré des résultats différents avec une diminution significative de champignons (abondance relative) et une transition vers des bactéries gram-positives et plus précisément vers les actinomycètes (Frey et al., 2008). Ces différents résultats sont probablement dus à la différence de T ° de base des différents sites où ces études ont été faites, d'autant plus que la variation saisonnière de la T ° est très importante sur l'un des deux sites. Par conséquent, l'effet sur les champignons (abondance relative) peut être dû à des différences de disponibilité de l'eau, et à l'effet du changement de T° sur cette disponibilité, entre les sols étudiés.

Gradient d'élévation

Comme pour le gradient de température, l'effet, pour les macroorganismes, du gradient d'élévation a été bien étudié par les écologistes depuis l'avènement de la biogéographie. Des pics de diversité de végétaux et de richesse phylogénétique sont observables à mi-altitude (500 à 1200m selon la latitude) (Austrheim et al., 2002, Kromer et al. 2005, Tang et al., 2003). Ce motif uni modale de distribution non linéaire (hump-shape), qui a été décrit pour les végétaux le long d'un gradient d'élévation, est très différent de ce qui a été observé pour les micro-organismes, en particulier les bactéries. Il a été mesuré que la diversité et la richesse bactérienne diminuent de façon monotone de la plus basse à la plus haute altitude et qu'à tous les niveaux d'élévation les bactéries ont tendance à être plus phylogénétiquement « clustérisées » (Bryant et al. 2008). Seules les plantes ligneuses présentent un schéma similaire de distribution le long du gradient d'élévation, mais ont tendance à disparaître totalement après une certaine altitude (ce n'est pas le cas pour les bactéries). Pour les très hautes altitudes (entre 4000 à 6500m), où les plantes sont presque absentes, une étude a montré le même genre de modèle de diminution linéaire des plus bas aux plus élevés échantillons pour les AOA et les AOB (Zhang et

al., 2009). En revanche, la plus faible diversité observable en haute altitude, suggère que seulement quelques micro-organismes se sont adaptés à ces environnements extrêmes. Plusieurs axes de recherche doivent encore être étudiés pour une meilleure compréhension de la distribution spatiale des bactéries à travers un gradient d'élévation. Toutefois, on peut déjà considérer l'altitude comme une des variables importantes contrôlant la distribution des microorganismes dans les études à grande échelle en zone de montagne.

Précipitation

Pour les macro-organismes, les quantités de précipitations sont fortement corrélées à la diversité. De la forêt tropicale où l'on observe de fortes précipitations, aux zones désertiques où il n'y a presque pas de précipitations, nous pouvons mesurer une baisse constante dans la diversité et dans la densité d'individus. Pour les micro-organismes (bactéries et les archées), en revanche, les taux de précipitations ne sont pas en lien direct avec la diversité. Une étude, du désert du Néguev (<100 mm de pluie par an) à la forêt méditerranéenne (> 900 mm de pluie par an), n'a montré aucune diminution significative de la diversité en fonction du gradient de précipitations (Angel et al., 2009). Fait intéressant, ils ont observé statistiquement de fortes différences en termes de structure de communauté entre les différents types d'écosystèmes, ce qui suggère que les précipitations et la couverture végétale, n'ont pas diminué la diversité, mais ont favorisé des structures de communautés spécifiques, regroupant les zones arides, les zones semi-arides et méditerranéennes séparément (ACP). Une autre étude a montré que l'abondance relative des champignons était fortement corrélée à l'humidité du sol (précipitations), bien que la diversité n'était pas affectée (Frey et al., 1999).

Couverture végétale

Lors d'études à grandes échelles, certaines recherches ont porté sur l'influence de la couverture végétale sur la structure de la communauté de bactéries habitant les sols. Le couvert végétal explique une plus grande part de la variabilité que les indicateurs classiques physico-chimiques (Mitchell et al., 2008). Une comparaison de différents types de végétation dans la région subarctique a souligné l'importance de la couverture de végetaux sur la structure de la communauté (Chu et al., 2011). Ils n'ont observé que de faibles différences dans la diversité entre les sites, mais l'abondance relative des principaux phylums différait selon le type de végétation (lande sèche, bouleau, grand bouleau, et carex humide). Parallèlement à cette étude, dans un environnement froid, une autre étude réalisée sur des sols de régions tempérées à des sols de régions plus chaudes, a démontré des résultats similaires (Chan et al.2008). Ils ont décrit des structures de communautés distinctes en fonction de la couverture végétale. Les zones boisées étant dominées par les Acidobacteria (62% des séquences totales), les zones d'arbustes avaient une abondance plus faible d'Acidobacteria et étaient dominées par les Betaproteobacteria (31%), tandis que les zones de pâturages étaient "dominées" par les Alphaproteobacteria (19%) et Bacteriodetes (16%). Les micro-organismes jouent un rôle important dans la rhizosphère lors de l'échange de ressources entre la plante et le sol. Les communautés vivant dans la rhizosphère sont différentes des communautés que l'on retrouve sous les sols nus. Dans un champ de maïs, la variation de l'abondance relative des taxons semblait être significative entre les échantillons collectés sous les sols nus, comparativement à ceux collectés dans la rhizosphère, sous les plants de maïs (Peiffer et al. 2013). Le sol nu étant l'hôte d'une plus grande diversité (principalement Acidobacteria, Bacteroidetes et Proteobacteria) par rapport à celle de la rhizosphère du même champ (principalement

Gammaproteobacteria des genres Pseudomonas et Lysobacter) (García-Salamanque et al. 2014). La spécialisation ou l'accroissement de populations spécifiques modifient la structure de la communauté vers une plus grande abondance des bactéries associées à des fonctions relatives à l'intéraction avec les végétaux. Dans une autre étude, une corrélation entre la rotation des cultures et la structure de la communauté bactérienne a été observée, bien que cet écart était inférieur à celui dû à la variation saisonnière (Maul et al. 2014).

Les paramètres physico-chimiques

Comme la plupart des paramètres physico-chimiques du sol ont une distribution hétérogène à micro-échelle, ces paramètres sont difficiles à inclure dans l'analyse spatiale à grande échelle de la distribution bactérienne dans les sols. Bien que les relations entre ces paramètres et la distribution des bactéries soit difficile à observer, certaines études ont proposé des liens entre pH (Shen et al. 2013, Fierer et al. 2012, Griffiths et al. 2012) et structure de la communauté du sol, plus précisément l'abondance des Acidobacteria. Ces relations ne sont pas nécessairement causales, mais pourraient refléter le rôle de la couverture de végétation qui est elle-même corrélée avec le pH du sol (Binkley et Fisher 2012). L'utilisation des terres (superficie agricole) peut également intégrer des variables telles que l'abondance de nutriments (C, N, P) et de structure du sol qui ont été décrites comme étant structurante de la communauté bactérienne du sol (Leff et al. 2015).

Les variables à méso-échelle

La méso-échelle (de l'échelle métrique à kilométrique) est l'échelle à laquelle la plupart des études environnementales sont faites. Même sous une surface relativement homogène, la répartition spatiale des bactéries dans le sol a été décrite comme très hétérogène et dépendante de variables qui se produisent à des échelles

spatiales plus petites (Franklin et al. 2003). Une meilleure compréhension des paramètres opérant à ces échelles aiderait à déterminer l'effet des pratiques d'utilisation des sols (principalement agricoles) sur la structure de la communauté bactérienne et ses rôles dans le fonctionnement des écosystèmes et de la santé des plantes.

Distribution des variables physico-chimiques et des nutriments

Des études ont porté sur des fonctions spécifiques associées à des bactéries ou archées pour trouver les variables influencant la distribution spatiale microbienne à l'échelle du champ. Il a été démontré que la répartition spatiale des bactéries AOB et archées AOA arborait un motif géographique spécifique à l'échelle de l'hectare (Wessén et al., 2011). L'abondance des AOB a été positivement corrélée avec le pH du sol, l'humidité du sol, le carbone organique total et l'azote total, tandis que l'abondance des AOA était corrélée négativement avec le pH du sol et la teneur en argile. D'autres recherches sur la distribution des communautés dénitrifiantes dans un champ de prairies a montré un modèle de corrélation spatiale à l'échelle de 6 à 16m (Phillipot et al., 2009). Elles démontrent également la corrélation spatiale entre la présence de bouses bovines (abondantes sur le terrain) et les communautés dénitrifiantes. Bien que certains éléments de preuve du lien potentiel entre distribution des paramètres édaphiques du sol et distribution spatiale des microorganismes, ces variable prises séparément peuvent difficilement être associées à une réponse dans la communauté microbienne. Les caractéristiques du sol semblent plus influentes que l'utilisation des terres (pratique) sur la structure de la communauté comme le montre une analyse de « variance partioning » (Thomson et al. 2015), bien que la petite taille du jeu de données limite la possibilité d'en tirer des conclusions reproductibles.

La distance spatiale

Une question cruciale est de savoir quelle distance (entre les échantillons) peut expliquer la diversité génétique (et génomique) entre deux échantillons. Les distances spatiales affectent l'hétérogénéité génétique à des échelles où l'hétérogénéité environnementale survient. La similitude du génome-entier devrait être fonction de la distance spatiale et de l'hétérogénéité de l'environnement, mais des paramètres environnementaux non mesurés (ou non mesurable), la variabilité spatiale et les biais d'échantillonnage, rendent difficile la démonstration de cette assertion. Une étude a estimé que l'effet pur de la distance spatiale n'explique que 2% de la variation génétique totale. Comme de multiples facteurs sont à l'origine de la distribution spatiale des bactéries dans les sols, et que ces facteurs multiples se produisent à de multiples échelles, nous pouvons détecter la considérable structure spatiale. Cependant, il est encore difficile d'isoler un seul facteur pour mesurer son impact sur la distribution spatiale des bactéries (Franklin et al., 2009).

Variables à micro-échelle

A la micro-échelle, l'échelle des bactéries, les sols sont très hétérogènes. Une grande diversité bactérienne et une grande hétérogénéité au sein de petits échantillons de sols (1 cm^3) ont été démontrés (Vogel et al., 2003), soutenant l'hypothèse que les paramètres physico-chimiques sont à l'origine de la distribution spatiale des bactéries (à cette échelle ce sont les variables qui affichent également une grande variabilité et hétérogénéité). Peu d'études ont porté sur les variables de la distribution bactérienne à micro-échelle dans les sols. Même si nous savons que les bactéries ne sont pas distribuées au hasard à l'échelle microscopique (Ranjard & Richaume 2001), nous sommes confrontés à une limite, puisque nous perdons toutes les informations spatiales à micro-échelle lors du traitement des échantillons

pour des études génomiques. Malgré ces limites, certains paramètres de la diversité spatiale ont été mis en évidence à la micro-échelle.

La taille des agrégats

La granulométrie à micro-échelle des sols peut varier considérablement et présente une très grande hétérogénéité spatiale. Les agrégats peuvent varier de 2 µm à plusieurs mm de diamètre. Certaines études ont tenté de séparer les agrégats par leur taille et de comparer les communautés bactériennes qui peuplent les différentes fractions granulométriques. Ces fractions ont différentes compositions organiques et minérales. Les micro-agrégats avec la plus faible teneur en carbone organique ont été associés à de fortes abondances relatives de bactéries de l'ordre des rubrobacteriales (Davinic et al. 2012). D'autre part, dans les mêmes sols, les macro-agrégats, avec une teneur en carbone organique plus élevée, ont été associés à une dominance d'Actinobacteria (à l'exclusion des rubrobacteriales). En comparant différents sols, les micro-agrégats ont été associés à de forte concentration de gemmatimonadetes, rubrobacteriales et de bactéries de la lignée des Alphaproteobacteria, alors que les macro-agrégats était dominés par les Acidobacteria (Mummey et al., 2006).

Variabilité chimique

Des études sur les sols ont montré que les différents types de matières organiques du sol peuvent être corelés avec les différentes fractions de taille de particules. Des analyses mi-infrarouges ont montré des caractéristiques distinctes spectrales pour les fractions plus grandes, matière organique particulaire, par rapport aux fractions de la tailles des silts et des argiles, soutenant l'idée d'une grande variabilité chimique

à micro-échelle (Calderon et al., 2011). La profondeur du sol joue également un rôle avec la diminution de la teneur en carbone organique en fonction de la profondeur dans de nombreux sols. Récemment, l'utilisation de « nano-scale secondary ion mass spectrometry » (NanoSIMS) sur des sols a fourni des preuves de zones de distribution et d'accumulation hétérogènes pour le Fe et le Mn (Rennert et al. 2014) suggérant que les bactéries associées pourraient suivre le même schéma de distribution. En combinant les NanoSIMS et les techniques « FISH », il a également été démontré la possibilité d'examiner la distribution microbienne et l'activité métabolique microbienne dans des échantillons environnementaux en les liants avec la distribution spatiale des variables chimiques (Chen et al. 2015).

Eau, pores, et réseau d'air

A l'échelle des bactéries, leur densité spatiale n'est pas très élevée, leurdéplacement est souvent contrôlé par la présence et le mouvement de l'eau. Les sols sont structurés avec un réseau d'air autour des agrégats et des pores remplis d'eau ou d'air et une pellicule d'eau peut recouvrir certains agrégats (Young et al., 2008). A l'aide de la microscopie, nous pouvons quantifier les bactéries à l'intérieur de ces pores et sur ces pellicules d'eau. La taille même de ces pores semble influencer la répartition des bactéries, les Phylum Actinobacteria et Firmicutes ont été identifiés comme étant plus abondant dans les grands pores par rapport à de petits pores (Kravchenko et al. 2014).

Résumé chapitre 1 suite

Quelle est la taille d'un métagénome de sol ?

Deux termes, sol et métagénome, doivent être définis avant de pouvoir aborder la question de la taille du métagénome du sol. Les sols sont des environnements très

complexes en termes de paramètres biotiques et abiotiques. Les Pédologues et géologues ont défini les grands groupes de sols en fonction des caractéristiques physico-chimiques. De la macro à la micro-échelle, les microorganismes du sol semblent être spatialement organisés. Le sol est également un environnement changeant et toute méthode qui prend une mesure rapide de la communauté microbienne du sol dépend de l'aspect temporel des processus microbiens. Le terme «métagénome» représente le total de tous les génomes présents dans un environnement, un écosystème, ou un échantillon. Ainsi, l'évaluation de l'ensemble du génome du sol n'est pas possible actuellement (biais lors de l'extraction de l'ADN, la PCR, le séquençage, l'analyse de données ...), le résultat d'une analyse métagénomique est un jeu de données métagénomiques. La taille d'un métagénome du sol est la somme de la taille de tous les génomes Archaea, procaryotes et eucaryotes présents dans notre échantillon. Ici, nous allons nous concentrer sur le métagénome bactérien et traiter du nombre d'organismes dans l'écosystème par rapport aux séquences d'un ensemble de jeux de données métagénomiques.

Chapitre 2 résumé

Overcoming micro-scale heterogeneity in a centimeters scale study of diesel contaminated soils

Les bactéries ont colonisé toutes les niches de la planète. Plus particulièrement, les sols sont l'habitat privélégié de la plus grande biodiversité terrestre, la faune microbienne. Cette très grande diversité des bactéries et leur relative ubiquité font en sorte qu'il est difficile d'identifier les variables qui contrôlent la distribution spatiale des bactéries habitant le sol. Comme les bactéries du sol jouent d'importants rôlesdans les principaux cycles biogéochimiques globaux, il est important de mieux

connaître les variables qui peuvent influencer la composition bactérienne des sols. Comme il fut démontré dans le chapitre 1, les variables contrôlant la distribution spatiale des bactéries surviennent à différentes échelles, de la micro échelle (l'échelle des bactéries) à la macro échelle (échelle régionale ou continentale). Les bactéries interagissent en premier lieu avec leurs environnements immédiats, les paramètres du sol variant à l'échelle micro auront donc une influence direct sur la composition bactérienne du sol. Il est toutefois très difficile d'analyser ces variations à micro échelle pour déterminer à de plus grandes échelles (significatives pour des études environnementales) la distribution des bactéries peuplant les sols. Afin de valider notre hypothèse que des variables influant à une échelle spatiale plus grande peuvent intégrer une partie de la variabilité imputée aux paramètres qui agissent à une échelle plus petite (micro), nous proposons dans cette étude d'induire un changement majeur sur une demie carotte de sol en ajoutant une quantité significative de diesel (200ml). Notre hypothèse spécifique est que la nouvelle variable « contamination au diesel » sera suffisamment influente pour structurer la composition bactérienne malgré la haute hétérogénéité à micro-échelle. Les résultats démontrent un transfert vers des taxons connus pour leurs capacités à utiliser le diesel comme source de carbone. La variable induite dans le système étant suffisamment importante pour surpasser les variations contrôlées par des variablesinfluantes à micro-échelle. Nous démontrons donc ici l'importance d'identifier des variables ayant une dimension spatiale compatible avec l'échelle spatiale de l'échantillonnage et la possibilité d'influer sur la composition bactérienne des sols à plusieurs échelles simultanément. D'autre part, les échantillons prélevés à différentes profondeurs démontrent la très grande variation spatiale dans l'axe vertical, pouvant surpasser la variation horizontale.

Résumé chapitre 3

Geographic Information System (GIS) Analyses of Large-Scale Spatial Soil Bacterial Diversity

Les sols sont probablement, pour la fraction microbiologique, les écosystèmes les plus riches en biodiversité. Malgré de considérables efforts de séquençage d'ADN et de rARN pour de nombreux types de sol, beaucoup restent à explorer pour comprendre comment ces communautés bactériennes sont structurées, étendent leurs interactions et leurs rôles dans le fonctionnement des écosystèmes. La distribution spatiale des bactéries habitant le sol est hautement hétérogène, à différentes échelles, mais demeure peu connu. Des études ont toutefois démontré l'existence de liens entre la distribution spatiale des micro-organismes avec la distribution spatiale de paramètres physico-chimiques du sol (*e.g.*, relation entre le pH du sol et l'abondance relative des *Acidobacter*). Dans ce projet, nous amenons l'hypothèse que l'hétérogénéité de la composition des communautés bactériennes du sol apparaît à la même échelle que les propriétés environnementales du sol. Pour la première fois dans le cadre d'une étude de terrain à grande échelle, une combinaison d'analyses par puces phylogénétiques, d'analyses physico-chimiques, et d'analyses spatiales à grande échelle en utilisant des systèmes d'informations géographiques (SIG), ont été utilisées pour étudier la distribution spatiale des bactéries dans le sol, afin de comprendre la relation entre la composition bactérienne du sol et les paramètres environnementaux du sol. Les analyses multivariées des résultats d'analyses phylogénétiques sur puces et des analyses physico-chimiques n'ont laissé voir aucune évidence de relations spécifiques entre les caractéristiques du sol et leurs compositions bactériennes, tout particulièrement pour les niveaux taxonomiques supérieurs. Par contre, avec les analyses spatiales par systèmes

d'informations géographiques, nous avons pu mettre en évidence la complexité des paramètres du sol qui contrôlent la structure des communautés microbiennes à l'échelle de très large régions avec l'exemple du pourcentage de couvert forestier versus le pH et les effets sur le phylum Acidobacteria.

Résume chapitre 4

Spatial and Temporal Analysis of Bacterial Diversity in Lake Sediment

Les sédiments sont, avec les sols, les écosystèmes les plus divers sur la planète. La distribution spatiale des communautés bactériennes habitant les sédiments est hautement hétérogène à différentes échelles spatiales et cette variabilité spatiale a été très peu explorée. Dans ce chapitre nous amenons l'hypothèse que l'hétérogénéité spatiale des communautés bactériennes varie à la même échelle que l'hétérogénéité spatiale des propriétés chimiques des sédiments. Nous nous intéresserons à la diversité bactérienne des sédiments à macro-échelle (Km). Selon la littérature, les variables physico-chimiques qui peuvent avoir une incidence sur la distribution spatiale des bactéries à cette échelle sont la couverture du sol des bassin-versants, le climat, le pH et la salinité. Pour tester cette hypothèse nous avons examiné la distribution spatiale des bactéries dans les sédiments du Lac Chilika (Inde) le deuxième plus grand lac d'eaux saumâtres au monde, le plus grand d'Asie. Soixante-douze échantillons (24 stations, 3 saisons – Hiver, mousson et été) de sédiments du lac Chilika furent analyser par pyroséquensage du gène 16S rRNA (région V4-V6). L'analyse de la couverture de surface a été réalisée avec des images satellites (Landsat) et des modèles d'élévation digitale à l'aide des logiciels GRASS et QuantumGIS. Un large spectre d'analyses physico-chimiques (e.g. pH, turbidity, salinity, conductivity, nitrate) a étéfurent réalisé sur l'eau et les sédiments pour chaque station d'échantillonnage et pour chaque saison. Apres un ouragan qui

dévasta la région en 2013, des échantillons supplémentaires ont étéfurent collectés afin de mesurer l'influence de phénomènes climatiques extrêmes (tempête tropicale) sur la distribution spatiale des bactéries dans les sédiments. Les résultats des analyses de l'rRNA 16S et des paramètres physico-chimiques interpolés géographiquement démontrent clairement une relation spatiale entre la distribution de paramètres physico-chimiques (salinité), géomorphologiques (drainage, fermes aquatiques) et la distribution des communautés microbiennes habitant le sédiment.

Conclusion

L'importance de la compréhension de la distribution spatiale des bactéries dans l'espace et des paramètres qui induisent cette répartition spatiale est critique. Les micro-organismes sont la première forme de vie sur la terre et tout ce qui est venu après est dû à leur activité et leur implication dans les grands cycles biogéochimiques. Dans un monde en mutation, l'impact des activités humaines sur la vie a été bien étudié, mais nousen connaissons toujours très peu sur les micro-organismes. Dans cette étude, nous avons identifié certaines tendances, à différentes échelles spatiales, qui tendent à confirmer notre hypothèse générale "La distribution spatiale des bactéries dans les sols et les sédiments est contrôlée par des paramètres physico-chimiques, la couverture terrestre et l'utilisation des terres, et que la variabilité spatiale se produit aux mêmes échelles spatiales que ces variables ".

Comme présenté dans le premier chapitre, la taille du métagénome du sol dépend de ce que nous considérons comme une unité de sol, la taille de l'échantillon et la profondeur du séquençage. Bien que rien ne ressemble plus à un métagénome de sol qu'un autre métagénome de sol, entre les différents écosystèmes, des

différences significatives en terme de structure de la communauté peuvent être observées et ce, à des échelles multiples. A l'échelle microscopique, la taille de l'environnement bactérien, quelques paramètres ont été identifiés pour comprendre comment les bactéries sont organisées dans l'espace. Certaines études ont identifié des micro-niches, mais pas nécessairement de structures de communautés associées à ces niches. Nous savons que les bactéries favorisent les pores remplis d'air ou recouverts d'un film d'eau et différents types de communautés habitent ces différentes niches. La Taille des agrégats a beaucoup été étudiée grâce à la possibilité de recueillir plusieurs agrégats de la même taille pour avoir suffisamment de matière pour faire un séquençage. Les corrélations entre la variabilité spatiale des paramètres physico-chimiques (principalement le type de matière organique) et de la taille des agrégats semblent structurer les communautés habitant ces agrégats. À l'échelle du champ (méso-échelle), les variables qui influencent la structure de la communauté bactérienne sont principalement la couverture végétale et la distribution des nutriments. Encore une fois, et plus particulièrement dans les champs agricoles, seules les variables ayant une hétérogénéité spatiale à l'échelle du champ peuvent être utilisées pour trouver des corrélations avec l'abondance relative de certains taxons. À une plus grande échelle (de régionale à l'échelle mondiale), quelques variables ont été identifiées qui contrôlent la structure de la communauté. Les variables arborant un modèle de distribution (patch, croissant, décroissant) à macro-échelle peuvent être utilisées pour modéliser la distribution spatiale des bactéries.

La question «Es-ce qu'une variable avec une répartition spatiale à plus grande échelle (cm) peut atténuer l'hétérogénéité à micro-échelle?" a été abordé dans le chapitre 2. Pour répondre à cette question, nous avons induit une modification

chimique importante sur une carotte de sol en ajoutant du diesel sur la moitié de la carotte. Après 14 jours d'incubation de la partie contaminées et non contaminées de la carotte, nous avons pu tester notre hypothèse «si l'ampleur de la perturbation correspond à l'échelle de l'échantillonnage, nous devrions observer un changement dans les organismes qui sont adaptés au nouvel état et « d'effacer » une partie de l'hétérogénéité à micro-échelle ». Nous avons observé une augmentation significative des taxons connus pour être présents dans les sols contaminés par des hydrocarbures. En outre, nous devons considérer les sols comme un environnement en 3 dimensions car nous avons trouvé des différences verticales supérieures à celles horizontales en termes de composition de la communauté, et l'effet de contamination a été atténué dans les échantillons les plus profonds.

Pour les études environnementales, nous devons considérer des surfaces beaucoup plus importantes afin de mesurer l'impact de changements dans l'environnement sur les cycles biogéochimiques globaux. Mais, puisque les paramètres physico-chimiques sont difficiles à modéliser à grande échelle, la question «Est-ce que certains bio-indicateurs à grande échelle peuvent intégrer des groupes de variables pour modéliser la distribution des bactéries pour toute une région ?" prend tout son sens. En couplant puces phylogénétiques, analyses physiques et chimiques, techniques d'analyses multivariées et systèmes d'informations géographiques à grandes échelles (SIG), nous avons comparé l'influence des paramètres physico-chimiques, seuls ou en groupe, avec une analyse de la couverture de surface. L'exemple des effets du pH (variable unique) sur les abondances relatives d'Acidobacteria (variable unique) par rapport à la variable pourcentage de forêt dans l'air de drainage (variables intégratives) a soutenu notre hypothèse que les variables qui intègrent un nombre holistique de données physico-chimiques (par exemple,

forêts) peuvent être de meilleurs indicateurs de la structure de la communauté que les données physico-chimiques prisent seules. Cette approche a également démontré la faisabilité d'utiliser des outils de SIG, les images satellites et les DEM pour l'analyse spatiale de la distribution bactérienne dans les sols. Ces outils permettent de tenir compte du transport de la matière (ruissellement) et ainsi incorporer une dimension temporelle.

Enfin dans le chapitre 4, nous avons présenté une étude de terrain sur les sédiments du lac Chilika (la première description temporelle à grande échelle de la distribution des populations bactériennes dans les sédiments d'un lac d'eaux saumâtres) afin de vérifier in situ l'hypothèse "Les variables avec des gradients géographiques à l'échelle de l'échantillonnage doivent être de puissants indicateurs de la distribution spatiale des bactéries et de la structure des communautés ". Avec une stratégie d'échantillonnage couvrant le gradient de salinité dans le lac, nous avons trouvé une forte corrélation entre la salinité et l'abondance de Proteobacteria et plus spécifiquement Gammaproteobacteria (la classe la plus abondante). L'utilisation d'outils d'analyses spatiales nous permet aussi d'identifier des « hot spots » de la diversité où nous avons des changements saisonniers de salinité. Avec des approches statistiques simples, les variables salinité et saison, étaient prises séparément et ne constituaient pas des indicateurs significatifs de la diversité microbienne. En outre, des analyses d'images satellites (aquaculture) et la modélisation hydrologique (flux hydrauliques) ont permis d'identifier des « hot spots » pour certains membres des Gammaproteobacteria connus pour être des pathogènes des poissons associés à des activités d'élevage de poissons.

Dans cette étude, nous avons souligné l'importance de considérer plusieurs échelles pour comprendre la répartition spatiale des bactéries dans les sols et les sédiments. En intégrant des données de gènes 16S ARNr et les données physico-chimiques dans un système SIG, nous avons pu nous concentrer sur les variables ayant un modèle de distribution spatiale compatible avec la zone des échantillons que nous avons analysé. Pour aller plus loin, en superposant plusieurs couches de différentes variables agissant sur différentes échelles spatiales dans un système SIG, nous pouvons approcher une représentation plus globale de la distribution spatiale des bactéries et de modéliser l'impact des changements sur l'environnement sur leurs activités.

Introduction

The spatial distribution of different microbial species and their functions in soils is still poorly understood in spite of considerable work on visualizing and testing different soil samples. Some of the difficulty is due to the technical and cost limitations of exploring all the microorganisms at different spatial scales. On the other hand, macro-organisms have been well studied in term of how they are distribute in space and time, and what are the variables that drive that spatial distribution. Ecological principles have been derived in some cases from this macroorganism research, but their application to microorganisms is not always apparent due in part to species definition and horizontal gene transfer. Bacteria have existed and evolved on earth over billions of years. During that period, and considering that their evolution rate is higher than macroorganisms due to their high reproduction rates (multiplication or scissiparity) and to other adaptation strategies like horizontal gene transfer, they have evolved up to 10 million species (Sogin *et al.* 2006). To access this large diverse genetic reservoir, we face the problem of the large number of low abundant, but almost always present microorganisms ("low abundance biosphere"). In soils, the low relative abundance difficulty is compounded by the number of individual "species" found in a gram of soil. Some estimates suggest that "species" numbers are as high as 10^7 per gram of soil (Gans *et al.* 2005). With the long history of bacteria on earth, they have had time to colonize every niche in soil all over the world. Wind transport, water transport and motility do not have to move bacteria more than 1 cm per year on average for bacteria to travel a total of 35 000 km (3.5 billion years ago), which is nearly the circumference of earth (40 075 km). Even continental boundaries cannot limit the dispersion since bacteria were already there long before the super continent

Pangaea 290 million years broke up even if not considering transport processes such as wind and water dispersion. So questions concerning whether everythin is everywhere have important temporal parameters. Yet, in soil, there is considerable quantities of uninhabited soil at the micron-scale and so clearly not everything is everywhere at that scale, but there exists scales, where there are sufficient numbers of bacteria (see comment above) and functions to drive global biogeochemical process (N cycling, carbon cycling, pollution cycling…), enhance or degrade plant health, and influence climate change. If this scale (or probably different scales for different processes) helps determine the contribution of bacteria inhabiting the soil, we need to understand the key variables that affect their presence and their activity as a function of space and time.

In the first chapter of this thesis, we will discuss the different variables that have been suggested as drivers of the soil bacterial community. There are multiple drivers proposed and they occur at different spatial scales. At the bacterial scale, the micro-scale, micro habitat, physical (e.g., aggregates, pore, water and air network) and chemical drivers are controlling the distribution of bacteria (Grundmann *et al.* 2004, Franklin & Mills 2003, Young *et al.* 2008). These physical and chemical characteristics are difficult to extrapolate to larger scales; scales that are more relevant for environmental studies (from field-scale to global scales), as we destroy most micro-spatial coherence by analyzing samples at the mg of soil scale. The meso scale (field-scale, meters to kilometers) studies have identified physico-chemical distribution, plant cover and nutrient availability as drivers of microorganisms, as associated with functions via assumed structure/function relationships (Fierer *et al.* 2009, Berg & Smalla 2009). At an even larger scale

(regional, continental, global), temperature gradient, elevational gradient, meteorological variables, geological substrate, land cover, pH, and human disturbance have been proposed as drivers for microbial community structure (Lauber *et al.* 2009, Fierer *et al.* 2012, Griffiths *et al.* 2011, Ge *et al.* 2008). In the second part of chapter 1, we will try to answer the question “what is the size of a soil metagenome”. Recently the rapid development of molecular biology methods has opened up access to more depth of sequencing, and therefore, hopefully better depth of analysis of microbial communities through metagenomic approaches (Delmont *et al.* 2012).

In the chapter two, we tried to answer the question “Can a variable (diesel contamination) at a larger scale (cm) overcome micro-scale heterogeneity?” We induced a change on half of a core of soil by adding 200ml of diesel to see if the community structure shifts to overcome the micro scale heterogeneity. Our hypothesis was that if the scale of the perturbation corresponds to the scale of the sampling, we should observe a shift in microorganisms that are able to use the substrate “diesel”. Those questions are important for spatial analysis at larger scale. For example do we have to consider the time scale when we look at variables like land cover? It’s also critical to have sampling strategies consistent with the scales of the variables influencing the community structure.

In chapter 3, we tried to answer the question “Do some large scale bioindicators integrate groups of variables to model the distribution of bacteria for an entire region”. We choose an entire region of northern France (250 km by 250 km). A large set of physico-chemical parameters have been produced for each sampling point. By doing

spatial analyses with GIS software, we compared the statistical approach with physico-chemical data and land cover to analyze the community structure. Our hypothesis was that integrating physico-chemical data can help correlations with community structure but that geographic information systems provide critical data about dynamic environmental processes such as run-off that can influence community structure.

Finally, in chapter four, we tested the hypothesis that “variables with geographical gradients at the sample size scale of the sampling are drivers of community structure”. We decided to do the study with sediment of Chilika Lake (India), the second world largest brackish lake. The lake has a salinity gradient from fresh water to marine water. Using a set of physico-chemical data, 16S rRNA gene analysis and GIS analysis we identified drivers (e.g., salinity, hydraulic flow) of microbial diversity and distribution.

Bibliography

Berg, G., & Smalla, K. (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS microbiology ecology*, *68*(1), 1-13.

Delmont, T. O., Simonet, P., & Vogel, T. M. (2012). Describing microbial communities and performing global comparisons in the 'omic era. *The ISME journal*, *6*(9), 1625-1628.

Fierer, N., Carney, K. M., Horner-Devine, M. C., & Megonigal, J. P. (2009). The biogeography of ammonia-oxidizing bacterial communities in soil. *Microbial ecology*, *58*(2), 435-445.

Fierer, N., Strickland, M. S., Liptzin, D., Bradford, M. A., & Cleveland, C. C. (2009). Global patterns in belowground communities. *Ecology Letters*, *12*(11), 1238-1249.

Franklin, R. B., & Mills, A. L. (2003). Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *FEMS Microbiology Ecology*, *44*(3), 335-346.

Gans, J., Wolinsky, M., & Dunbar, J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science*,*309*(5739), 1387-1390.

Ge, Y., Zhang, J. B., Zhang, L. M., Yang, M., & He, J. Z. (2008). Long-term fertilization regimes affect bacterial community structure and diversity of an agricultural soil in northern China. *Journal of Soils and Sediments*, *8*(1), 43-50.

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of British soils. *Environmental microbiology*, *13*(6), 1642-1654.

Grundmann, G. L. (2004). Spatial scales of soil bacterial diversity–the size of a clone. *FEMS Microbiology Ecology*, *48*(2), 119-127.

Lauber, C. L., Hamady, M., Knight, R., & Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and environmental microbiology*,*75*(15), 5111-5120.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., ... & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proceedings of the National Academy of Sciences*, *103*(32), 12115-12120.

Young, I. M., Crawford, J. W., Nunan, N., Otten, W., & Spiers, A. (2008). Microbial distribution in soils: physics and scaling. *Advances in agronomy*,*100*, 81-121.

Part 1

What are the main drivers of soil bacterial community structure and at what scale do they occur

Introduction

Life, in the form of microorganisms, has been present on earth for over 3 billion year. This long and varied history has led to the evolution of bacteria that resulted in their colonization of every niche of our planet. The extremely high diversity of bacteria compared to macro-organisms is probably due in part to that long period of time and that their growth and adaptation rates are much faster. Since the beginning of microbial ecology, we have been asking the question “is everything everywhere?”. Bass-Becking in 1934 was the first to propose an answer that “everything is everywhere but the environment selects”. Soils are always highly diverse, but no other ecosystem has a more similar functional fingerprint than another soil (Delmont et al. 2011). In order to understand how soil micro-organisms are organized in soil and who are where (and doing what), we have to determine the drivers of the spatial distribution of bacteria. Understanding the link between physico-chemical parameters and the community structure of soils will provide the basis for modeling the global spatial distribution of bacteria in soils and hopefully the changing environment.

In this review, we will describe the different drivers of the spatial distribution of bacteria in soils that have been identified. We will deduce the different scales at which these drivers occur. Different variables have to be considered with field studies depending on the scale we are working on and the sampling density we use. Everything happens at the micro-scale, the scale of the bacteria, but it is difficult to

model (or at least validate the model) a large area from extrapolated micro-scale data due to the high heterogeneity of soil characteristics. Some variables have been identified as reasonable indicators of community structure from the centimeters scale to the continental scale. Some of those variables might be useful as indicators.

Larger-scale indicators

Temperature

The first global scale parameter that we have to focus on is the temperature. Most macro-organism have limited dispersion that is partially related to their acceptable temperature ranges. In the macro-organism world, probably only humans have been able to colonize every range of the temperature gradient, although that is mainly due to our capacity to change our environment (air conditioning, heating…) and to artificially protect ourselves with clothes. In the microbiology world, links between temperature and spatial distribution of micro-organisms is not so clear. Micro-organism are more diverse in the cold North (Neufeld *et al.* 2005) where their community structure is not so different from what was observed in warmer soil (Zhang *et al.* 2005, Chu *et al.* 2010). Recently, considerable bacterial diversity has been observed in arctic snow (Larose *et al.* 2010, Macario *et al.* 2014). On the other hand, micro-organisms are also present and diverse in the warmest environment on earth even where no macro-organisms can survive like nearby the ocean ridge (Huber *et al.* 2003).

According to the problematic of climate changes, some studies have focus on the consequences of an increase of temperature (2 to 5°C) in soils and the effect on the

community structure. A shift to higher relative abundance of fungi compare to bacteria was observed after 2 to 5 years of artificially increased temperature (Castro *et al.* 2010, Zhang *et al.* 2005). A longer study (12 years of 5°C increase above ambient T°), have shown different result with a significantly decrease of fungi relative abundance and a shift toward gram-positives bacteria and Actinomycetes (Frey *et al.* 2008). Those different results are probably due to the difference in the T° of the different sites where studies were made, seasonal variation of T° is quite important and the effect on fungi relative abundance can be due to soil water availability differences.

Elevation

Like for the temperature (T°) gradient, the effect of elevation has been well studied for macro-organism by ecologists since the beginning of the biogeography science. Plant diversity has maximum richness and phylogenetic diversity at mid-elevation (500 to 1200m depending on the latitude) (Austrheim *et al.* 2002, Kromer *et al.* 2005, Tang *et al.* 2003). These unimodal patterns and hump-shape patterns that have been describe for plant across an elevational gradient is a lot different from what we observed for micro-organisms, specifically bacteria. Bacterial diversity and richness have been described to decrease monotonically from the lowest to the highest elevation and at all elevations, bacteria tend to be more phylogenetically clustered (Bryant *et al.* 2008). Only woody-plants exhibit a similar pattern of distribution across the elevational gradient, but tend to totally disappear at a certain elevation (which is not the case for bacteria). For extremely high elevations (between 4000 to 6500m), where plants are almost absent, a study has shown the same kind of monotonical pattern of bacterial and archaeal ammonia oxidizers from lowest to highest elevations

(Zhang *et al.* 2009), which implies that only some micro-organisms have adapted to this extreme environment. More research needs to be done on the spatial distribution of bacteria across elevational gradients, but elevation can already be consider as an important driver of microbial community structure and diversity for large scale studies.

Precipitation

For macro-organisms, precipitation is strongly correlated to their diversity. From the high precipitation tropical forest to the almost no precipitation desert area, we can measure a constant drop in the diversity and the numbers of individuals. For micro-organisms (bacteria and archaea), precipitation rate are not link with the diversity. A study has shown no significant decrease in the diversity from the Negev Desert (< 100mm rain per year) to the Mediterranean Forest (>900mm rain per year) as a function of the precipitation gradient (Angel *et al.* 2009). Interestingly they observed statistically strong differences in terms of community structure between the different types of ecosystems, suggesting that precipitation and vegetation cover, did not decrease the diversity but did drive the community structure as the arid, semi-arid and Mediterranean climatic zones clustered separately. Another study has shown that the relative abundance of fungi is strongly correlated to soil moisture (precipitation), although diversity was unaffected (Frey *et al.* 1999).

Vegetation cover

In a large scale study, some research has focused on the influence of specific vegetation on the community structure of bacteria inhabiting the soil. The vegetation

cover explains more of the variability then classic physico-chemical indicators (Mitchell *et al.* 2008). Comparing different types of vegetation in the sub-arctic region highlighted the importance of the vegetation cover on the community structure (Chu *et al.* 2011). They found no major differences in the diversity between sites, but the relative abundance of the main phyla differed significantly and consistently according to the vegetation type (dry heath, birch hummock, tall birch, and wet sedge). In parallel to that study in a cold environment, another one from temperate to warm environments had similar results (Chan *et al.*2008). They described distinct community structures based on the vegetation cover. Forested areas were dominated by Acidobacteria (62% of the total sequences), shrub-land had lower abundance of Acidobacteria and were dominated by Betaproteobacteria (31%), and pasture was “dominated” by Alphaproteobacteria (19%) and Bacteriodetes (16%). Micro-organisms play an important role in the rhizosphere during the exchange of resources between the plant and its soil environment. The communities inhabiting soils in the rhizosphere are different than communities under bare soil. In a corn field, variation of relative abundances of taxa appeared to be significant between bulk soils and soils in the corn plant rhizosphere (Peiffer *et al.* 2013). The bulk soil from the corn field was home to more diversity (mainly Acidobacteria, Bacteroidetes and Proteobacteria) compared to that in the rhizosphere of the same field (mainly Gammaproteobacteria of the genera Pseudomonas and Lysobacter) (García-Salamanca *et al.* 2014). In another study, they correlated crop rotation and bacteria community structure, although this variance was less than that due to season variation (Maul *et al.* 2014).

Physico-chemical parameters

As most of the soil physico-chemical parameters did not correlate with large-scale geographical characteristics, these parameters are difficult to include in large-scale spatial analysis of bacterial distribution in soils. Although any relationship between these parameters and bacteria distribution is difficult to observe, some studies proposed links between pH (Shen *et al.* 2013, Fierer *et al.* 2012, Griffiths *et al.* 2012) and soil community structure, specifically the abundance of Acidobacteria. These relationships are not necessarily causal ones, but might reflect the role of vegetation cover that is itself correlated with soil pH (Binkley & Fisher 2012). Land-use (agricultural area) can also integrate variables like nutrient abundance (C,N, P) that have been described to affect the soil bacterial community structure (Leff *et al.* 2015).

Meso (field, landscape) scale variables

The field scale is the scale at which most environmental studies are made. Even in a relatively homogenous field, the spatial distribution of bacteria in soils has been described as highly heterogeneous and dependent on variables that occur at that spatial scale or smaller (Franklin *et al.* 2003). A better understanding of the drivers operating at these scales would help to determine the effect of the soil use practices (mainly agricultural) on the bacterial community structure and their role in ecosystem functioning.

Physico-chemical and nutrient parameters

Studies have focused on specific groups of bacteria or archaea to find drivers of microbial spatial distribution at the field-scale. The spatial distribution of the

ammonia-oxidizing bacteria (AOB) and ammonia-oxidizing archaea (AOA) were shown to have a specific geographical pattern at the hectare scale (Wessén *et al.* 2011). The abundance of AOB was positively correlated with soil pH, soil moisture, total organic carbon and total nitrogen while AOA abundance was negatively correlated with soil pH and clay content. Other research on the distribution of the denitrifying community in a grassland field showed a pattern of spatial-correlation at the scale of 6 to 16m (Phillipot *et al.* 2009). They also demonstrate the spatial correlation between the presence of cattle (abundant on the field) and the denitrifying communities. Although some evidence of the potential link between edaphic condition and microorganism spatial distribution in soils exists, a single variable cannot be easily associated to a response in the microbial community. Soil characteristics seem more important than land use (practice) on the community structure as shown by a variance partioning analysis (Thomson *et al.* 2015), although the importance of the dataset in drawing conclusions needs to be assessed.

Spatial distance

One critical question is at what distance scale (between samples) can explain the genetic (and genomic) diversity between two samples. The spatial distances affect genetic heterogeneity at scales where environmental heterogeneity occurs. The whole-genome similarity should be function of the spatial distance and environmental heterogeneity, but unmeasured environmental parameter, spatial variability and sampling effect make it difficult to demonstrate. A study has estimated that the pure effect of spatial distance explained only 2% of the total genetic variation. As multiple soil factors are driving the spatial distribution of bacteria in soil, and as those multiple factor occurred at multiple scales, we can detect considerable spatial structure but it

still difficult to isolate one factor to measure their impact on the spatial distribution of bacteria (Franklin *et al.* 2009).

Micro-scale variables

At the micro-scale, the scale of bacteria, soils are highly heterogeneous. We observed a high bacterial diversity and heterogeneity within small samples of soils (Vogel *et al*. 2003), supporting the hypothesis that physico-chemical parameters are driving the spatial distribution of bacteria. Few studies have focused on the micro-scale drivers of bacterial distribution in soils. Even if we know that bacteria are not randomly distributed at microscale (Ranjard & Richaume 2001), we face a limit while we lose all micro-scale spatial information by processing the samples for genomic studies. Despite these limits, some drivers of the spatial diversity have been highlighted at the micro-scale.

Size of aggregates

The micro-scale granulometry of soils can vary considerably in spatial heterogeneity. Aggregates can range from 2 µm to several mm in diameter. Some studies have attempted to separate aggregates by their size and compare the bacterial communities that inhabit those different granulometric fraction. These fractions have different organic and minerals composition. The micro-aggregates with the lowest organic carbon content were associated with high relative abundance of bacteria from the order *rubrobacteriales* (Davinic *et al*. 2012). On the other hand, macro-aggregates with the highest organic carbon content from the same soils were associated with a dominance of *Actinobacteria*, excluding the order *rubrobacteriales*.

Comparing different soils, micro-aggregates were dominated by *Gemmatimonadetes*, *rubrobacteriales* and lineage of *alphaproteobacteria* while macro-aggregate were dominated by *acidobacteria* (Mummey *et al.* 2006)

Chemical differences

Soil studies have shown that different types of soil organic matter correlated with different size fractions. Mid-infrared analysis has shown distinct spectral features for the light fraction, particulate organic matter, silt-sized and clay sized fractions supporting the idea of chemical differences (Calderon *et al*. 2011). Soil depth also plays a role with decreasing organic carbon content with depth in many soils. Recently, the use of nano-scale secondary ion mass spectrometry (NanoSIMS) on soils has provided some evidence of heterogeneous distribution and accumulation zones for Fe and Mn (Rennert *et al*. 2014) suggesting that associated bacteria might follow the same pattern of distribution. NanoSIMS in combination with FISH has also demonstrated the possibility of examinating microbial distribution and microbial metabolic activity in environmental samples link with chemical distribution (Chen *et al.* 2015).

Water, pores and air network

At the bacteria-size-scale, the coverage of soils by is not very high and their movement is often controlled by the presence and movement of water. Soils are structured with a network of air around aggregates and pores filled with water or air and water film (Young *et al.* 2008). We can quantify the bacteria inside these pores and on these water films with microscopy. Actinobacteria and Firmicutes groups were

observed to be more abundant in large pores compared to small pores (Kravchenko *et al*. 2014).

Bibliography

Angel, R., Soares, M. I. M., Ungar, E. D., & Gillor, O. (2010). Biogeography of soil archaea and bacteria along a steep precipitation gradient. *The ISME journal*, *4*(4), 553-563.

Austrheim, G. (2002). Plant diversity patterns in semi-natural grasslands along an elevational gradient in southern Norway. *Plant Ecology*, *161*(2), 193-205.

Bryant, J. A., Lamanna, C., Morlon, H., Kerkhoff, A. J., Enquist, B. J., & Green, J. L. (2008). Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences*, *105*(Supplement 1), 11505-11511.

Binkley, D., & Fisher, R. (2012). *Ecology and management of forest soils*. John Wiley & Sons.

Calderón, F. J., Reeves, J. B., Collins, H. P., & Paul, E. A. (2011). Chemical differences in soil organic matter fractions determined by diffuse-reflectance mid-infrared spectroscopy. *Soil Science Society of America Journal*, *75*(2), 568-579.

Castro, H. F., Classen, A. T., Austin, E. E., Norby, R. J., & Schadt, C. W. (2010). Soil microbial community responses to multiple experimental climate change drivers. *Applied and environmental microbiology*, *76*(4), 999-1007.

Chan, O. C., Casper, P., Sha, L. Q., Feng, Z. L., Fu, Y., Yang, X. D. & Zou, X. M. (2008). Vegetation cover of forest, shrub and pasture strongly influences soil bacterial community structure as revealed by 16S rRNA gene T-RFLP analysis. *FEMS microbiology ecology*, *64*(3), 449-458.

Chen, C., Bai, Y. H., Liang, J. S., & Yuan, L. J. (2015). [Application of FISH-NanoSIMS technique in environmental microbial ecology study]. *Huan jing ke xue= Huanjing kexue/[bian ji, Zhongguo ke xue yuan huan jing ke xue wei yuan hui" Huan jing ke xue" bian ji wei yuan hui.]*, *36*(1), 244-251.

Chu, H., Fierer, N., Lauber, C. L., Caporaso, J. G., Knight, R., & Grogan, P. (2010). Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environmental Microbiology*, *12*(11), 2998-3006.

Chu, H., Neufeld, J. D., Walker, V. K., & Grogan, P. (2011). The influence of vegetation type on the dominant soil bacteria, archaea, and fungi in a low Arctic tundra landscape. *Soil Science Society of America Journal*, *75*(5), 1756-1765.

Davinic, M., Fultz, L. M., Acosta-Martinez, V., Calderón, F. J., Cox, S. B., Dowd, S. E., ... & Moore-Kucera, J. (2012). Pyrosequencing and mid-infrared spectroscopy reveal distinct aggregate stratification of soil bacterial communities and organic matter composition. *Soil Biology and Biochemistry*,*46*, 63-72.

Delmont, T. O., Robe, P., Cecillon, S., Clark, I. M., Constancias, F., Simonet, P., ... & Vogel, T. M. (2011). Accessing the soil metagenome for studies of microbial diversity. *Applied and Environmental Microbiology*, *77*(4), 1315-1324.

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., ... & Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences*, *109*(52), 21390-21395.

Franklin, R. B., & Mills, A. L. (2003). Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *FEMS Microbiology Ecology*, *44*(3), 335-346.

Franklin, R. B., & Mills, A. L. (2009). Importance of spatially structured environmental heterogeneity in controlling microbial community composition at small spatial scales in an agricultural field. *Soil Biology and Biochemistry*,*41*(9), 1833-1840.

Frey, S. D., Elliott, E. T., & Paustian, K. (1999). Bacterial and fungal abundance and biomass in conventional and no-tillage agroecosystems along two climatic gradients. *Soil Biology and Biochemistry*, *31*(4), 573-585.

Frey, S. D., Drijber, R., Smith, H., & Melillo, J. (2008). Microbial biomass, functional capacity, and community structure after 12 years of soil warming.*Soil Biology and Biochemistry*, *40*(11), 2904-2907.

García-Salamanca, A., Molina-Henares, M. A., Dillewijn, P., Solano, J., Pizarro-Tobías, P., Roca, A., ... & Ramos, J. L. (2013). Bacterial diversity in the rhizosphere of maize and the surrounding carbonate-rich bulk soil. *Microbial biotechnology*, *6*(1), 36-44.

Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of British soils. *Environmental microbiology*, *13*(6), 1642-1654.

Huber, J. A., Butterfield, D. A., & Baross, J. A. (2003). Bacterial diversity in a subseafloor habitat following a deep-sea volcanic eruption. *FEMS Microbiology Ecology*, *43*(3), 393-409.

Kravchenko, A. N., Negassa, W. C., Guber, A. K., Hildebrandt, B., Marsh, T. L., & Rivers, M. L. (2014). Intra-aggregate pore structure influences phylogenetic composition of bacterial community in macroaggregates. *Soil Science Society of America Journal*, *78*(6), 1924-1939.

Krömer, T., Kessler, M., Robbert Gradstein, S., & Acebey, A. (2005). Diversity patterns of vascular epiphytes along an elevational gradient in the Andes.*Journal of Biogeography*, *32*(10), 1799-1809.

Leff, J. W., Jones, S. E., Prober, S. M., Barberán, A., Borer, E. T., Firn, J. L., ... & Fierer, N. (2015). Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academy of Sciences*, *112*(35), 10967-10972.

Larose, C., Berger, S., Ferrari, C., Navarro, E., Dommergue, A., Schneider, D., & Vogel, T. M. (2010). Microbial sequences retrieved from environmental samples from seasonal Arctic snow and meltwater from Svalbard, Norway.*Extremophiles*, *14*(2), 205-212.

Maccario, L., Vogel, T. M., & Larose, C. (2014). Potential drivers of microbial community structure and function in Arctic spring snow. *Frontiers in microbiology*, *5*.

Maul, J. E., Buyer, J. S., Lehman, R. M., Culman, S., Blackwood, C. B., Roberts, D. P., ... & Teasdale, J. R. (2014). Microbial community structure and abundance in the rhizosphere and bulk soil of a tomato cropping system that includes cover crops. *Applied Soil Ecology*, *77*, 42-50.

Mitchell, R. J., Hester, A. J., Campbell, C. D., Chapman, S. J., Cameron, C. M., Hewison, R. L., & Potts, J. M. (2012). Explaining the variation in the soil microbial community: do vegetation composition and soil chemistry explain the same or different parts of the microbial variation?. *Plant and soil*, *351*(1-2), 355-362.

Mummey, D., Holben, W., Six, J., & Stahl, P. (2006). Spatial stratification of soil bacterial populations in aggregates of diverse soils. *Microbial Ecology*,*51*(3), 404-411.

Neufeld, J. D., & Mohn, W. W. (2005). Unexpectedly high bacterial diversity in arctic tundra relative to boreal forest soils, revealed by serial analysis of ribosomal sequence tags. *Applied and Environmental Microbiology*, *71*(10), 5710-5718.

Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., ... & Ley, R. E. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences*,*110*(16), 6548-6553.

Philippot, L., Čuhel, J., Saby, N., Chèneby, D., Chroňáková, A., Bru, D., ... & Šimek, M. (2009). Mapping field-scale spatial patterns of size and activity of the denitrifier community. *Environmental microbiology*, *11*(6), 1518-1526.

Ranjard, L., & Richaume, A. (2001). Quantitative and qualitative microscale distribution of bacteria in soil. *Research in Microbiology*, *152*(8), 707-716.

Rennert, T., Händel, M., Höschen, C., Lugmeier, J., Steffens, M., & Totsche, K. U. (2014). A NanoSIMS study on the distribution of soil organic matter, iron and manganese in a nodule from a Stagnosol. *European Journal of Soil Science*, *65*(5), 684-692.

Shen, C., Xiong, J., Zhang, H., Feng, Y., Lin, X., Li, X., ... & Chu, H. (2013). Soil pH drives the spatial distribution of bacterial communities along elevation on Changbai Mountain. *Soil Biology and Biochemistry*, *57*, 204-211.

Tang, Z., & Fang, J. (2003). A review on the elevational patterns of plant species diversity. *Biodiversity Science*, *12*(1), 20-28.

Thomson, B. C., Tisserant, E., Plassart, P., Uroz, S., Griffiths, R. I., Hannula, S. E., ... & Lemanceau, P. (2015). Soil conditions and land use intensification effects on soil microbial communities across a range of European field sites.*Soil Biology and Biochemistry*, *88*, 403-413.

Vogel, J., Normand, P., Thioulouse, J., Nesme, X., & Grundmann, G. L. (2003). Relationship between spatial and genetic distance in Agrobacterium spp. in 1 cubic centimeter of soil. *Applied and environmental microbiology*,*69*(3), 1482-1487.

Wessén, E., Söderström, M., Stenberg, M., Bru, D., Hellman, M., Welsh, A., ... & Hallin, S. (2011). Spatial distribution of ammonia-oxidizing bacteria and archaea across a 44-hectare farm related to ecosystem functioning. *The ISME journal*, *5*(7), 1213-1225.

Young, I. M., Crawford, J. W., Nunan, N., Otten, W., & Spiers, A. (2008). Microbial distribution in soils: physics and scaling. *Advances in agronomy*,*100*, 81-121.

Zhang, W., Parker, K. M., Luo, Y., Wan, S., Wallace, L. L., & Hu, S. (2005). Soil microbial responses to experimental warming and clipping in a tallgrass prairie. *Global Change Biology*, *11*(2), 266-277

Zhang, L. M., Wang, M., Prosser, J. I., Zheng, Y. M., & He, J. Z. (2009). Altitude ammonia-oxidizing bacteria and archaea in soils of Mount Everest.*FEMS Microbiology Ecology*, *70*(2), 208-217.

Chapter 1

Part 2: What is the Size of a Soil Metagenome

Abstract

Two terms, soil and metagenome, need to be defined in this context in order to evaluate the size of the soil metagenome. Soils are highly complex environments in terms of biotic and abiotic parameters. Pedologist and soil scientist have defined major groups of soil depending on physico-chemical characteristics. From the macro- to the micro-scale, soil microorganisms seem to be spatially organized. Soil is also a changing environment and any method that takes a rapid measurement of the soil microbial community is dependent on the temporal aspect of microbial processes. The term “metagenome” is the total of all the genomes present in an environment, ecosystem, or sample. Since evaluating the entire genome of soil is not currently possible (bias during DNA extraction, PCR, sequencing, data analysis…), the result of the metagenomic approach is a metagenomic dataset. The size of a soil metagenome is the sum of the size of all Archaea, Prokaryote and Eukaryote genomes present in our sample. Here, we will focus on the bacterial metagenome and address the number of organisms in the ecosystem versus the sequences in a metagenomic dataset.

Introduction

Microorganisms play an important role in soil ecosystems by contributing to biogenic transformation of carbon, nitrogen and phosphorus (Paul and Clark, 1989). Indirectly microorganisms also contribute to plant nutrition (Timonen et al. 1996), plant health (Filion et al. 1999), soil fertility (O'Donnell et al. 2001) and soil structure. Soils are also the largest reservoir of genetic material and we are now observing a new *gold rush* in term of research for new antibiotics or novel therapeutic compounds isolated from soil bacteria (Raaijmakers *et al.* 1997). Recently, several molecular biology methods (including DNA-DNA reassociation, DNA cloning and sequencing and fingerprinting approaches) have been used to estimate bacterial diversity in soil (Frostegard et al 1993, Kowalchuk et al 1998, Muyzer et al 1993, Torsvik et al 1990, Robe *et al.* 2003). The high concentration of bacteria and archea in soils has been estimated between 10^7 to 10^9 per gram and the number of "species" or taxa for the same gram of soil between 10^3 to 10^7(Ovreas *&* Torsvik*.* 1998, Gans *et al.* 2005, Mende *et al.* 2012). Soils are the largest reservoir in quantity and diversity of microorganisms. The cellular production rate for all prokaryotes on earth is estimated at 1.7×10^{30} cells/yr and the total microorganisms in all soils is calculated at around 2.6×10^{29} (Withman *et al.* 1998). This reservoir of genetic material is still poorly described and metagenomic approaches can help us to access the vast amount of information held within soil bacterial genomes. In the recent years, several studies have been done to understand what are the drivers of the bacterial composition of soils and their spatial distribution. Metagenomic studies on soils have been conduct at a wide range of scale from continental scale (Fierer et al. 2006, Griffiths et al. 2011) to micro scale (Grundmann et al. 2004, Vos et al. 2013, Remenant et al. 2009).

The limitation we are now facing is to sample representative soils in order to characterize them and to limit the bias that can occur during the post sampling analysis. The 10^6 to 10^8 taxa existing in one gram of soil (Torsvik *et al.* 2002, Gans *et al.* 2005) are the result of 3.5 billion years of bacterial adaptation and evolution. With a higher rate of "evolution" than any other life form, they are able to live in any type of soil. These complex community structures are in perpetual transformation due to large potential of adaptation like gene transfer and high reproduction rates.

With metagenomic studies, we seek to link the genetic potential of the biome from a given ecosystem and its function in order to be able to understand, predict and influence this function. The question arises whether our current approaches in soil ecological studies are suitable to describe the whole functional potential of soils.

Metagenome and soil samples

In order to determine the size of a soil metagenome we first have to understand what is a metagenome and what we define as a soil. Traditionally studies in soil science followed a pedological approach in which soil are classified and described as they occur or an edaphological approach where soil is assessed as a plant production environment (Pal et al. 2012). Since microorganisms were found to substantially influence the soil ecosystem, microbial ecologists added a new angle of view to this complex environment. Soil, the most diverse ecosystem we know, covers around 87% of the landmass of this planet (Latham et al. 2014). When we talk about soil we often describe the pedosphere, which is material layer build up by interaction

between the lithosphere, atmosphere, biosphere and hydrosphere. This thin layer serves as a reactor of microbial conversion processes in all major nutrient cycles (Nielsen et al. 2011; Ollivier et al. 2011) and as an essential reservoir for plants. Many factors have to be accounted for when it comes to soil sampling for an ecological study, to ensure representative results and relate them to descriptive soil parameters (Perkins et al. 2013) Sampling soil to describe and characterize the inhabiting communities usually included a sieving step to get rid of plant roots, gravel and higher biota, because of issues during DNA isolation.

Diversity, distribution and abundance of higher biota in soil

Higher biota communities in soil are, beside their taxonomy, distinguished by their size. Terms describing faunal groups in literature are macrofauna (bodysize from 1 mm to several centimeters) and meio- or mesofauna (bodysize < 1mm) (Bik et al. 2012). Alternatively to laborious identification by hand, size spectra of higher organisms from different soil ecosystems are used describe changes in these communities and link them to function (Turnbull et al. 2014). According to these studies bodysizes between 150 µm (Rotifera) up to 30 centimeters and more (Coleoptera larvae, Chilopoda etc.) are found.

Between 273 and 503 macrofaunal organisms were found in the topsoil (15 cm) of a forest in Borneo (Hasegawa et al. 2014) per m^2. Scaling these numbers to the amount of soil sampled and extracted in metagenomic studies would result in 0.03 macrofaunal organisms per cubic centimeter of this soil. For mesofaunal communities in average 3 individuals could be found per cm^3 of topsoil. In average

11 groups were identified for macro-, 40 groups for mesofaunal members. From another study one can calculate 0.015 macrofaunal individuals per cm^3, whereby 16 taxa were identified. Much higher numbers were obtained for nematodes, rotifers and tardigrades in the topsoil (10 cm) of a Swedish scots pine forest (Sohlenius 1979). Between 110 and 630 nematodes per cm^3 of topsoil were counted over the period of a year, fluctuations over time which have to be considered in estimating the size of the soil metagenome. Numbers also changed significantly between different soil layers in the first 10 cm of this soil (539, 518 and 38 nematodes g^{-1} dry soil for S-layer, FH-layer and mineral soil respectively). Summarized this shows that soil fauna critically influencing function can be present in high abundances and diversities and should be considered in metagenomic surveys.

Excluding the faunal soil community from the metagenome does not just result in overseeing these members of the ecosystem. Microbial communities inhabiting the gut of higher biota were shown to have an impact on function and could be clearly distinguished from surrounding soil communities (Egert et al. 2003). From this point of view higher biota might serve as a mobile habitat for specialized microbial communities, probably playing key roles in nutrient conversion processes in soils.

Nevertheless, incorporating faunal communities in metagenomic studies faces problems. One is the size of the organisms what suggests a revision of soil sampling methods for these studies. Another might be sequencing depth problems due to the higher amount of DNA of these organisms as the one of microorganisms, thus “suppress” the number of sequences acquired from microorganisms in the same sequencing event.

First attempts profiling these communities with molecular tools were already done (Wu et al. 2009). Also a linear relationship between the biomass of the soil inhabiting beetle *Fosomia candida* and results from a qPCR approach (Hou et al. 2014) was shown suggesting the possibility of describing a soil ecosystem by metagenomics including higher biota and their gut inhabiting microorganisms.

Soils microbial metagenome

A metagenome is per definition the sum of genomes of all organisms from a given community in a given environment at a given time (Foster 2012). If we exclude plants and higher biota form our studies, the question arises if we are describing the soil metagenome with our current approaches at all, although of course there are good reasons why we perform this simplifications like problems with DNA extraction and sequencing depth. Since plants and higher biota actively influence soil ecosystems (Baker 2013, Frisli 2013) we have to improve our tools to access metagenomic data in order to be able to include all participants forming the soil ecosystem in our studies.

Still, at the current stage of available methods, due to the bias appearing during DNA extraction, PCR, sequencing and data analyzing steps, it is impossible to access to the entire metagenome of a soil sample. More accurately we should talk about metagenomics datasets, subsets of the metagenome that we've been able to capture during a sequencing event (Rodriguez and Konstantinidis 2014). Another important point is that the resulting sequences obtained by metagenomic sequencing have to

be annotated using incomplete databases. As around 1% of soils bacteria are known to be easily cultivable, we have a lack of accuracy for some of those non cultivable taxa. This is especially true when it comes to annotation of eukaryotic sequences from metagenomic datasets (Bik 2012). From Nematodes, who account for 80-90% of all metazoans and 1 million species are estimated, only 4% are formally known yet.

What is representative as soil for microbiologist?

Volume

In order to do metagenomics analysis on soil we have to work with quantity of soil around 1 gram. Often, the gram of soil used for the analysis, is a homogenate of several larger samples. That size of sample is in fact way too large to be used to understand the spatial distribution of bacteria in soils (Vos *et al.* 2013). Proportionally, it's like blending a whole forest, extracting the DNA and trying to understand the spatial distribution of plants in that landscape from the metagenomics data set produced. New sequencing technologies will soon allow us to sequence quantity of soils more proportional to the size of bacterial communities. But, even with smaller sample, it's almost impossible to find a sample that will be representative of a larger soil system because from the micro to macro scale, soils have very high heterogeneity of abiotic and biotic conditions (Beare et al, 1995, Ramette et al. 2007).

Deepness

Another limitation in the search of a representative sample of soils is the deepness of the soil sample, the variation in term of community structure is higher vertically then horizontally. For the same soil, the deepest sample of soils could have as much as 40% less diversity then the first 10 cm of soils (Eilers *et al.* 2012). We have to take in account the deeper fraction of soils, cause even if we can observe an exponential decrease of the density of bacteria in a depth gradient (Hartmann *et al.* 2009), two studies have estimated the proportion of the microbial biomass of the subsurface soil (deeper then 25 cm and 50 cm) to be 35 and 50% of all the soils bacteria (Fierer *et al.* 2003, Schutz *et al.* 2010). It's also important to consider the different horizons of soils.

What are the main taxa in soil

The dominant microorganisms in soils are bacteria. The proportion of bacteria can reach 95% of the microorganisms in prairies and desert, in temperate and boreal forest we observe a lower concentration of bacteria due to the abundance of fungi (Fierer *et al.* 2012). Archaea having there higher representation in desert soils but are often under 5% of total microorganisms (Bates *et al.* 2011). For bacteria the most represented phyla are the *Acidobacteria, Actinobacteria, Bacteriodetes, Proteobacteria* and *Verrucomicrobia.* Those phyla have been described to be relatively abundant in all type of biome from the hot desert to boreal forest (Janssen *et al.* 2006). Others phyla are also always represented, but with a lower relatives abundance, in all type of natural soils : *Chloroflexi, Cyanobacteria, Firmicutes* and *Gemmatimonadetes* (Fierer *et al.* 2012).

A matter of scale

Macro Scale

In a global scale not many significant correlations with the abiotic world have been elucidated. One exception is the correlation of the relative abundance of *Acidobacteria* with differences in soil pH (Fierer *et al.* 2006, Griffiths *et al.* 2009). These previous studies demonstrate correlations but do not necessarily show causation. In the case of pH variations, soil organic matter type also correlates with pH, with low pH forest soils at one extreme and high pH grasslands and desert soils at the other. But *Acidobacteria* cannot be consider as an homogenous phyla as some class of *Acidobacteria,* like Chloracidobacteria, are known to be favored by high pH soil (Jones *et al.* 2009) and have been measured in a global survey in higher proportion in desert compared to temperate forest soils (Fierer *et al.* 2012). In Term of diversity, pH has been highlighted as a strong predictor. Soils around a pH 7 are the one with the highest diversity, and acidic and basic soils have lower level of diversity (Griffiths *et al.* 2009). For one type of soils (grassland), a worldwide study (Prober *et al.* 2014) have also found a strong relation between plant beta-diversity and soils bacterial beta-diversity, suggesting that sites that were more distinct in the composition of their plant communities also shown more distinct soil microbial communities. In the same worldwide grassland study they didn't observed relation between alpha diversity of plants and microorganisms, suggesting that relationships between plant and soil microbial alpha diversity observed within sites (Milcu *et al.* 2013) may not persist when comparing sites at a global scale.

Another worldwide study on fungal communities analyzing 365 soil samples(Tedersoo et *al,* 2014) on all continents (except Antarctica) with the exact same procedures showed that fungal to plant richness ratio increased exponentially with distance from the equator, due to a fast decrease in plant diversity compared to fungal diversity with increasing latitude.

Meso Scale

At the meso-scale (field scale), principally vegetation cover (Uroz et al. 2010) and nutrient abundance (Philipot et al. 2009, Wessen et al. 2011) have been shown to correlate with phylogenetic and functional microbial diversity in soil. A single sample of soils from a field cannot be representatives of the whole field. Those studies showed the importance of blending together soils sample with plant or no plant cover and different range of nutrient abundance in order to have a representative sample for determining the field microbial metagenome. A recent study, Lauber *et al.* 2013, have also demonstrated the importance of temporal change in a field study. They show that the temporal variability (month of sampling) of alfa-diversity is higher than the variability between land use type.

Micro Scale

The micro-scale is poorly understood in term of spatial distribution of bacteria but it's the most relevant scale to look at in order to understand how bacteria are organized in space. Soils samples used for metagenomic studies (0.25 to 1 gram) are huge compare to the actual size of a micro-organism ($7x10^{-16}$ kg for Escherichia Coli) and sample processing destroy all spatial information on bacteria and their

resource (Holden *et al*. 2011). Several studies were done with the attempt to keep spatial information in the data by size-fractionation of soil (Bach et *al.*, 2014) (Ranjard et *al.,* 2000). Visualizing the soil at the scale of bacteria we can observe that bacteria are spatially organized in soils. Soil, at micro-scale, is a very complex environment with different size of aggregated particles (micro-aggregate mainly contain clay and humus particles and macro-aggregate mainly organic polymers, fungal hyphae and plant roots) (Six *et al*. 2004, Chenu & Consantino 2011) in a 3D network of water and air pores of different size that co-occur in close proximity (Young *et al.* 2008). The overall density of bacteria in soil is surprisingly low, with an estimation of soil surface covered by microorganisms at a mere 10^{-6}

(Young & Crawford, 2004) and a distance between colonies of hundreds of µm's (Grundmann *et al*. 2001). There is also a huge diversity of dormant-cells, outside of colonies, that play roles in the resilience of soil communities in case of environmental changes (Prosser *et al.* 2003).

Limitation to access the entire metagenome of a soil

If we can sample a representative sample of soil, the limitation to access the totality of that metagenome will be the limit of the technology and the database we use to compare our extracted sequence. Before sequencing, several steps have to be done and all of them will add bias to our metagenome construction. For example DNA or RNA extraction step will give similar but distinct result depending the commercial kit used (Mahmoudi *et al.* 2011, Delmont *et al.* 2011). Recent development of NGS tools allow us to access very large metagenomics dataset. Those environmental data

sets can be compared to nucleotide and protein databases like GO (Ashburner *et al.* 2000), COG (Tatusov *et al.* 2001), Pfam (Finn *et al.* 2010), NCBI (Sayers *et al*. 2011), SEED (Overbeek *et al*. 2005) and KEGG (Kanehisa *et al*. 2008). The statistical analysis can, then, be done using common statiscal software (R for example) but also with integrated platform like STAMP (Parks and Beiko 2010) or MG-RAST (Meyer *et al.* 2008). The result obtain will also depend on the database used and the statistical approach to interpret the produced metagenomic data set. The metagenome of one of the most studied soils, Park Grass Experiment at Rothamsted Research (Silvertown *et al.* 2006), have been use to construct a soil metagenome and to test reproducibility with different extraction protocol and different database (Delmont *et al.* 2012). They showed that even with duplicate samples (same extraction protocol, same database) we can observe significantly differences at the functional and taxonomic level. But comparing to other soils or to other biome (marine environment), the soils from Rothamsted were clustered together.

Conclusion

Soils are the most diverse ecosystem on earth. The important role play by micro-organisms inhabiting the soils have to be measure in order to know there part in global biogeochemical cycling and to model the consequences of climate change on their production. Several development in metagenomic allow us to go deeper in sequencing effort. But until now, we understand only a small part of how those bacterial communities in soils are organized in space and time. Some studies have describe drivers of the spatial distribution of bacteria at different scale level from the micro to macro scale, and by digging deeper under the top soil we discover that we

have to consider soils as 3D matrix showing a lot of variability along the depth gradient. The size of soil metagenome depend on the scale we working on, the deepness and a lot of pfysico-chemical parameters driving the presence or not of specific bacteria. The extremely high diversity of soils micro-organisms make every soil different from the others, but with development in bioinformatics tools we begin to be able to cluster soils together relating to environmental characteristics.

Bibliography

Bach and Hofmockel, 2014, Soil aggregate isolation method affects measures of intra-aggregate extracellular enzyme activity. *Soil Biology and Biochemistry*, *69*, 54-62.

Bates et al., 2011, Examining the global distribution of dominant archaeal populations in soil. *The ISME Journal* (2011) 5, 908–917.

Beare, M.H. *et al.,* 1995. A hierarchical approach to evaluating the significance of soil biodiversity to biogeochemical cycling. *Plant and Soil* 170: 5-22.

Bik, H.M. et al., 2012. Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in Ecology & Evolution*, 27(4), pp.233–243.

Chenu and Consantino, 2011, Microbial Regulation of Soil Structural Dynamic. *The Architecture and Biology of Soils*, pages 17-69

Delmont et al. Accessing the soil metagenome for studies of microbial diversity. *Appl. Environ. Microbiol.* 77, 1315–24 (2011).

Delmont et al., 2012, Structure, fluctuation and magnitude of a natural grassland soil metagenome. *The ISME journal,* 6(9), 1677-1687.

Egert, M. et al., 2003. Microbial community structure in midgut and hindgut of the humus-feeding larva of Pachnoda ephippiata (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology*, 69(11), pp.6659–6668.

Eilers, K. G., et al., 2012, Digging deeper to find unique microbial communities: The strong effect of depth on the structure of bacterial and archaeal communities in soil. *Soil Biol. Biochem.* 50, 58–65.

Fierer et al., 2003, Variations in microbial community composition through two soil depth profiles, *Soil Biology and Biochemistry*, Volume 35, Issue 1, January 2003, Pages 167–176

Fierer and Jackson,2006, The diversity and biogeography of soil bacterial communities. PNAS, vol. 103, 626–31 Fierer et al., 2012, Cross-biome metagenomic analyses of soil microbial communities and their functional attributes, *PNAS*, vol. 109 no. 52, 21390–21395

Fierer et al., 2012, Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *PNAS,* vol. 109 no. 52, 21390–21395

Filion et al., 1999, Direct interaction between the arbuscular mycorrhizal fungus *Glomus intraradices* and different rhizosphere microorganisms. *New Phytol.*, 141 (1999), pp. 525–533

Finn et al., 2010, The Pfam protein families database. *Nucleic acids research*, 38(suppl 1), D211-D222.

Foster et al., 2012, Measuring the microbiome: perspectives on advances in DNA-based techniques for exploring microbial life. *Brief Bioinform* 13(4):420–9.

Frisli et al., 2013, Estimation of metagenome size and structure in an experimental soil microbiota from low coverage next-generation sequence data, *Journal of Applied Microbiology*, Volume 114, Issue 1, pages 141–151

Frostegard et al., 1993, Phospholipid Fatty Acid Composition, Biomass, and Activity of Microbial Communities from Two Soil Types Experimentally Exposed to Different Heavy Metals. *Appl. Environ. Microbiol.* vol. 59 no. 113605-3617

Gans et al., 2005, Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science*, 309, 1387–1390.

Griffiths et al. 2011, The bacterial biogeography of British soils. Environ Microbiol13(6):1642–1654.

Grundmann et al., 2001, Spatial modeling of nitrifier microhabitats in soil. *Soil Science Society of America Journal*, 65(6), 1709-1716.

Grundmann, G. L. Spatial scales of soil bacterial diversity--the size of a clone. *FEMS Microbiol. Ecol.* 48, 119–27 (2004).

Hartmann et al., 2009, Bacterial, archaeal and eukaryal community structures throughout soil horizons of harvested and naturally disturbed forest stands. *Environmental Microbiology*, Volume 11, Issue 12, pages 3045–3062

Hasegawa, M. et al., 2014. The effects of reduced-impact logging practices on soil animal communities in the Deramakot Forest Reserve in Borneo. *Applied Soil Ecology*, 83, pp.13–21.

Holden et al., 2011, How do the microhabitats framed by soil structure impact soil bacteria and the processes that they regulate. *The architecture and biology of soils: life in inner space*. 118-149.

Hou, L.-B. et al., 2014. Biomass estimation of the terrestrial ecotoxicological species Folsomia candida (Collembola) using a real-time polymerase chain reaction. *Ecotoxicology and Environmental Safety*, 101(C), pp.59–63.

Janssen et al., 2006, Identifying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16S rRNA Genes. *Appl. Environ. Microbiol.* March 2006 vol. 72 no. 3 1719-1728

Jones et al., 2009, A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *The ISME Journal* 3, pp. 442–453

Kanehisa et al., 2008, KEGG for linking genomes to life and the environment. *Nucleic acids research*, 36(suppl 1), D480-D484.

Kowalchuk et al. 1998, Community analysis of ammonia-oxidising bacteria, in relation to oxygen availability in soils and root-oxygenated sediments, using PCR, DGGE and oligonucleotide probe hybridisation. *FEMS Microbiol. Ecol.*, 27 pp. 339–350

Latham, J. et al., 2014. FAO Global Land Cover (GLC-SHARE) Beta-Relaese 1.0 Database. pp.1–40.

Lauber et al., 2013, Temporal variability in soil microbial communities across land-use types. *The ISME Journal*, 7, 1641–1650

Mahmoudi et al., 2011, Comparison of commercial DNA extraction kits for isolation and purification of bacterial and eukaryotic DNA from PAH-contaminated soils. *Canadian journal of microbiology,* 57(8), 623-628.

Mende et al., 2012, Assessment of metagenomic assembly using simulated next generation sequencing data. *PLoS ONE*. 7(2):e31386.

Meyer et al., 2008, The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, 9(1), 386.

Muyzer et al., 1993, Phylogenetic relationships ofThiomicrospira species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Archives of Microbiology* Volume 164, Issue 3, pp 165-172

Nielsen, U.N., Ayres, E. & Wall, D.H., 2011. Soil biodiversity and carbon cycling: a review and synthesis of studies examining diversity–function relationships. *European Journal of Soil Science*, *62*(1), 105-116.

O'Donnell et al., 2001, Plants and fertilisers as drivers of change in microbial community structure andfunction in soils. *Plant Soil*, 232, pp. 135–145

Ollivier, J. et al., 2011. Nitrogen turnover in soil and global change. *FEMS Microbiology Ecology,* 78(1), pp.3–16.

Ollivier et al., 2011, Nitrogen turnover in soil and global change, *FEMS Microbiology Ecology*, , Pp. 3 - 16

Overbeek et al., 2005, The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic acids research, 33(17), 5691-5702.

Ovreas and Torsvik, 1998, Microbial diversity and community structure in two different agricultural soil communities *Microbiol. Ecol.*, 36, pp. 303–315

Parks and Beiko, 2010, Identifying biologically relevant differences between metagenomic communities. *Bioinformatics*, 26(6), 715-721.

Paul and Clark, 1989, Soil Microbiology and Biochemistry, Academic Press, London

Perkins *et al.,* 2013. Quick start guide to soil methods for ecologists. *Perspectives in Plant Ecology, Evolution and Systematics*, *15*(4), 237-244.

Philippot, L. et al., 2009, Spatial patterns of bacterial taxa in nature reflect ecological traits of deep branches of the 16S rRNA bacterial tree. *Environ. Microbiol.* 11, 3096–104.

Prober et al., 2014, Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecology Letters*, pages 85–95

Prosser et al., 2003, The influence of selection pressures on species diversity, functional gene diversity and activity of ammonia-oxidizing bacteria. In 'Genes in the Environment'. *The Fifteenth Special Symposium of the British Ecological Society.* (pp. 187-202). Blackwell Publishing.

Milcu et al., 2013, Functionally and phylogenetically diverse plant communities key to soil biota. *Esa journals*, Volume 94, Issue 8, pages 1878–1885

Raaijmakers et al., 1997, Frequency of Antibiotic-Producing Pseudomonas spp. in Natural Environments. *Appl. Environ. Microbiol.* vol. 63 no. 3 881-887

Ramette, A. Multivariate analyses in microbial ecology. *FEMS Microbiol. Ecol.* 62, 142–60 (2007).

Ranjard et al., A soil microscale study to reveal the heterogeneity of Hg(II) impact on indigenous bacteria by quantification of adapted phenotypes and analysis of community DNA fingerprints. Volume 31, Issue 2, Pp. 107 - 115

Remenant, B., et al., 2009, From the micro-scale to the habitat: Assessment of soil bacterial community structure as shown by soil structure directed sampling. *Soil Biol. Biochem*. 41, 29–36.

Robe et al., 2003, Extraction of DNA from soil, *European Journal of Soil Biology*, Pages 183–190

Rodriguez and Konstantinidis, 2014, Estimating coverage in metagenomic data sets and why it matters, *The ISME Journal* 8, 2349-2351

Sayers et al., 2011, Database resources of the national center for biotechnology information. *Nucleic acids research*, 39(suppl 1), D38-D51.

Schutz et al., 2010, Functional microbial community response to nutrient pulses by artificial groundwater recharge practice in surface soils and subsoils. FEMS Microbiology Ecology, 72, pp. 445–455

Silvertown et al., 2006, The Park Grass Experiment 1856–2006: its contribution to ecology.*Journal of Ecology*, 94(4), 801-814.

Six et al. A history of research on the link between (micro)aggregates, soil biota, and soil organic matter dynamics. September 2004, Pages 7–31

Sohlenius, B., 1979. A carbon budget for nematodes, rotifers and tardigrades in a Swedish coniferous forest soil. *Holarctic Ecology*, 2, pp.30–40.

Tatusov et al., 2001, The COG database: new developments in phylogenetic

classification of proteins from complete genomes. *Nucleic acids research*, 29(1), 22-28.

Tedersoo et al., 2014, Global diversity and geography of soil fungi. *Science*, Vol. 346 no. 6213

Timonen et al., 1996, Dynamics of phosphorus translocation in intact ectomycorrhizal systems: non-destructive monitoring using aβ-scanner. *FEMS Microbiol*. Rev. 19, 171-180

Torsvik et al., 1990, High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* Vol. 56 no. 3 782-787

Torsvik, V. & Øvreås, L. Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol*. 5, 240–5 (2002).

Turnbull, M.S., George, P.B.L. & Lindo, Z., 2014. Weighing in: Size spectra as a standard tool in soil community analyses. *Soil Biology and Biochemistry*, 68(C), pp.366–372.

Uroz, S., et al., 2010, Pyrosequencing reveals a contrasted bacterial diversity between oak rhizosphere and surrounding soil*. Environ. Microbiol.* Rep. 2, 281–8.

Vos, M., et al., 2013, Micro-scale determinants of bacterial diversity in soil. *FEMS Microbiol*. Rev. 37, 936–954.

Young et al. 2008, Microbial Distribution in Soils: Physics and Scaling., 2008, Pages 81–121

Young and Crawford, 2004, Interactions and Self-Organization in the Soil-Microbe Complex. *Science* Vol. 304 no. 5677 pp. 1634-1637

Wessén, E. et al., 2011, Spatial distribution of ammonia-oxidizing bacteria and archaea across a 44-hectare farm related to ecosystem functioning. *ISME J*. 5, 1213–25.

Wu, T. et al., 2009. Molecular profiling of soil animal diversity in natural ecosystems: Incongruence of molecular and morphological results. *Soil Biology and Biochemistry,* 41(4), pp.849–857.

Chapter 2

Overcoming micro-scale heterogeneity in a centimeters scale study of diesel contaminated soils

Résumé

Les bactéries ont colonisé tous les niches de la planète. Plus particulièrement, les sols servent d'habitat à la plus grande biodiversité terrestre, la faune microbienne. Cette très grande diversité des bactéries et leurs relatives ubiquité font en sorte qu'il est difficile d'identifier les variables qui contrôlent la distribution spatiale des bactéries habitant le sol. Comme les bactéries du sol jouent d'important rôle dans les principaux cycles biogéochimiques globaux, il est important de mieux connaître les variables qui peuvent influencer la composition bactérienne des sols. Comme il fut démontré dans le chapitre 1, les variables contrôlant la distribution spatiale des bactéries surviennent à différentes échelles, de la micro échelle (l'échelle des bactéries) à la macro échelle (échelle régionale ou continentale). Les bactérie interagissent en premier lieux avec leurs environnements immédiat, les paramètres du sols variant à l'échelle micro auront donc une influence direct sur la composition bactérienne du sol. Il est toutefois très difficile d'analyser ces variations à micro échelles pour déterminer à de plus grandes échelles (significatives pour des études environnementales) la distribution des bactéries peuplant les sols. Afin de valider notre hypothèse que des variables influant à une échelle spatiale plus grande peuvent intégrer une partie de la variabilité imputée aux paramètres qui agissent à une échelle plus petite (micro), nous proposons dans cette étude d'induire un changement majeur sur une demie carotte de sol en ajoutant une quantité significative de diesel. Notre hypothèse spécifique est que la nouvelle variable « contamination au diesel » sera suffisamment influente pour structurer la composition bactérienne malgré la haute hétérogénéité à micro-échelle. Les résultats démontrent un transfert vers des taxons connu pour leurs capacités à utiliser le diesel comme source de carbone. La variable induite dans le système étant

suffisamment important pour surpasser les variations contrôler par des invariable influant à micro-échelle. Nous démontrons donc ici l'importance d'identifier des variables ayant une dimension spatiale compatible avec l'échelle spatiale de l'échantillonnage et la possibilité d'influer sur la composition bactérienne des sol à plusieurs échelles simultanément. D'autre part, les échantillons prélevés a différentes profondeurs démontre la très grande variation spatiale dans l'axe vertical, pouvant surpassé la variation horizontale.

Abstract

The bacteria have colonized all the niches of the planet. Specifically, soils are home of the largest terrestrial biodiversity, microbial fauna. This great diversity of bacteria and their relative ubiquity make it difficult to idendified variables driving the spatial distribution of bacteria living in the soil. As soil bacteria play a significant role in the main global biogeochemical cycles, it is important to better understand the variables that can influence bacterial composion of soils. As shown in chapter 1, the variables driving the spatial distribution of bacteria occur at different spatial scales, from micro scale (the scale of bacteria) to the macro scale (regional or continental scale). The bacteria interact first with their immediate environment, soil parameters varying at micro scale will therefore have a direct influence on bacterial composition. It is very difficult to integrate micro scales variability in a larger scale study (significant for environmental studies) to quantified and qualified the distribution of bacteria populating the soils at macro-scale. To validate our hypothesis “variables affecting a larger spatial scale can overcome some of the variability attributed to the parameters which act on a smaller scale (micro)”, we propose in this study to induce a major change to a half soil core by adding a significant amount of diesel. Our specific hypothesis is that the new variable "diesel contamination" will be enough to structur the bacterial composition despite the high heterogeneity in micro-scale. The results demonstrate a shift towards taxa known for their ability to use diesel as a source of Carbon. The induced variable in the system is important enough to overcome variables driving microscale heterogeneity. We demonstrate here the importance of identifying variables with a spatial scale compatible with the spatial scale of sampling and the possibility to work at more environmentaly relevant scale. Furthermore,

samples taken at different depths demonstrated the very high spatial variation in the vertical axis may be more important then the horizontal variation.

Introduction

Bacteria are known to be almost everywhere on earth. They have colonized different type of habitat from extreme cold region like arctic snow (Larose *et al.* 2010) to extreme warm region like oceanic ridge (Kelley *et al.* 2002). Comparing those extreme environments we can observe very different bacterial communities composition. Those differences suggest that microbial world, like macro organism, have drivers that determine their spatial distribution. In soils, the extremely high diversity of bacteria, 10^3 to 10^7 species in a single gram of soil (Gans et al. 2005, Curtis et al. 2002, Tringe et al. 2005, Torsvik et al. 2002) and there relatively high ubiquity (Janssen et al. 2006) makes it difficult to find drivers of the bacterial spatial distribution of bacteria. As bacteria are involve in process regulating most of the nutrient cycling, plant growth and carbon storage required for life on earth; it's important to understand what are the drivers of the microbial world in soils and at what scale they occur. Specifically, in changing environment like the one we are facing today, the impact of those changes on bacterial spatial distribution can allow us to have a better understanding of the consequences on global biogeochemical cycling.

The major problem we face for the identifiying of the drivers of bacterial distribution in soils is the extreme heterogeneity of micro-environment. Due to size of the sample required for metagenomic or 16S analysis, and the impossibility to use the exact same sample for chemical, physical and biological composition analysis, it's almost impossible to extrapolate what we observed at micro-scale for a larger scale analysis.

We propose in that study to induce major changes in half of a core of soil by adding significant quantity of diesel. We hypothesis that the new variable “diesel contamination” will overcome the micro-scale heterogeneities and the variable “presence” or “absence” of diesel will be the main drivers of the spatial distribution of bacteria in terms of communities structures. In order to have global image of the changes induce by the contamination we also looked at different depth inside the cube of soil.

Material and method

Sampling

A cube of soils (30cm*30cm*30cm) was manually extracted with a shovel. The chosen soil was located in forest soil surrounded by grassland located nearby the campus of Ecole Centrale in Ecully. The extracted soils were then taken to the lab. The cube of soils was split in to half. We contaminated with diesel one of the two half. 200 ml of diesel was added on the top of the half cube in order to have a homogenous dispersion of the diesel(Figure 1).

The sub sampling of the two half of the cube was done 14 days after the contamination with diesel fuel. 24 top soil samples, following an equidistant matrix, were taken on each of the two half. 5 samples were taken at 6cm depth in each of the core, and one deeper (9cm) sample was taken on each half. Each of the sample was weighted at 250mg in order to have enough genetic material for the further analysis.

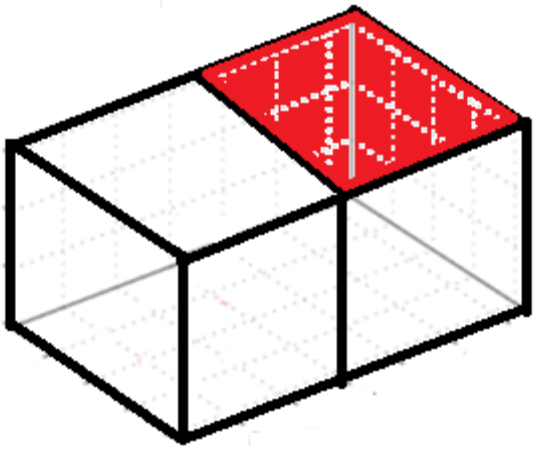


Figure 1: Cube of soils, diesel addition on the red part

DNA analysis

Total DNA were extracted from 0.25g of soil sample using PowerSoil®DNA Isolation Kit for soil DNA extraction (MO BIO, USA) as per the manufacturer's protocol. DNA was extracted was extracted 14 days after the addition of diesel (contaminated half core), the uncontaminated half core was extracted at the same moment. Quality and quantity of metagenomic DNA were verified by 1.2% of agarose gel (0.5XTAE) and Nanodrop (Epoch, BioTek Instruments, USA) analysis (A_{260}/A_{280}). DNA from a total of 60 samples were extracted for microbial diversity and community composition.

Total DNA extracted from the soil was amplified for 16S rRNA genes by PCR. The hyper variable (V4-V6) regions of the 16S rRNA genes was amplified from the total DNA extracted from each samples using bacterial primer pair 515F and 1061R (16S-0515F 5'-TGYCAGCMGCCGCGGTA-3' 16S-1061R 5'-TCACGRCACGAGCTGACG-3' ~560bp V4-V6 region). PCR was carried out with the modified primers containing an adapter and a barcode sequence (Schloss *et al*.,

2009). Paired end 454 pyrosequencing was performed on the GS-FLX 454 Titanium platform.

Spatial analysis

Interpolations (krigging) with GRASS software were performed for the main phylum distribution of the sample with and without diesel contamination, in order to find spatial changes in bacterial composition linked with the chemical and nutrient changes induce by the addition of diesel. To test the importance of the spatial distance we did statistical analysis of the spatial effect on the dissimilarity between samples (JUMP)

Results

16S rRNA gene analysis

Phylum Level

The result of the 16S rRNA gene analysis has shown a shift in the community structure after the diesel addition. At the phylum level we observed a dominance of the Proteobacteria in both contaminated and non-contaminated sample. The high diversity inside Proteobacteria may explain the fact that we were not able to observe stronger significant differences between samples. 34,8% to 50,1% (relative abundance) of all the sequence extracted were annotated as Proteobacteria.

The relative abundance of Actinobacteria was significantly higher after the diesel addition, reaching an average relative abundance of 26,5% compare to 16,2%

without diesel addition (Figure 2). That increase of Actinobacteria correspond to the result of other studies on hydrocarbon contaminated soils (Labbé *et al.* 2007, Chikere *et al.* 2009, Alvarez *et al.* 2008). That increase suggests an important contribution of the actinobacteria for the degradation of diesel in contaminated soils.

The half core contaminated with diesel, also shown a decrease in relative abundance of Acidobacteria, Verrucomicrobia and Bacteriodetes (Figure 2). Those results are similar to what have been observed in other oil contaminated sites (Saul *et al.* 2005). That decrease in relative abundance was only observable for the surface samples, the deeper sample (with or without diesel) had a much higher relative abundance of Bacteriodetes.

Relative abundance at the phylum level without diesel addition

Proteobacteria	0,40428298
Actinobacteria	0,161645
Acidobacteria	0,12625631
Verrucomicrobia	0,08468426
Bacteroidetes	0,09845891
Planctomycetes	0,05574433
Chloroflexi	0,02786645
Gemmatimonadetes	0,01628507
Firmicutes	0,00734054
Other	0,00460415
WS3	0,00272659
Armatimonadetes	0,00136487
Unclassified	0,00168903
Elusimicrobia	0,00142214
Nitrospirae	0,00103112
Chlorobi	0,0012575
Cyanobacteria	0,00061628
TM7	0,00035924
TM6	0,00040688
BRC1	0,0003715

Relative abundance at the phylum level with diesel addition

Proteobacteria	0,42213382
Actinobacteria	0,26460007
Acidobacteria	0,10674525
Planctomycetes	0,06465025
Verrucomicrobia	0,05192649
Bacteroidetes	0,03044589
Chloroflexi	0,02821943
Gemmatimonadetes	0,01483068
Firmicutes	0,00577091
Other	0,0032769
WS3	0,00143398
Unclassified	0,00107382
Armatimonadetes	0,00101731
Nitrospirae	0,00073445
Cyanobacteria	0,00057545
Elusimicrobia	0,00052722
Spirochaetes	0,00045087
TM7	0,00028844
Chlorobi	0,00024726
TM6	0,00023282

Figure 2: 20 most important (relative abundance) Phylum with and without diesel contamination

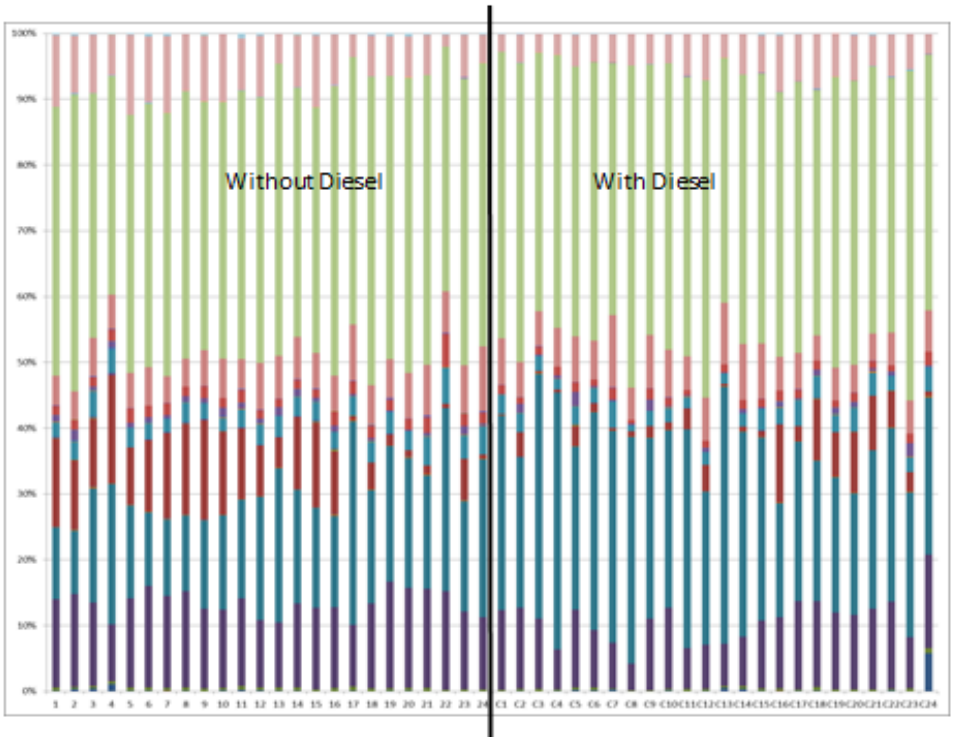


Figure 3: Distribution of the relative abundances of phylum in all the top soil samples

Order level

At the order level (Figure 4) we observed an important increase of Actinomycetales and Rhizobiales relatives abundance from the uncontaminated core to the contaminated one, respectively 11.13% to 22.05% and 12.68% to 17.33%. The increase of Actinomycetales is mainly due to the capacity of some species from that order to use diesel as a carbon substrate. Specifically a lot of species of Rhodococcus have been identified as degraders of diesel or other hydrocarbs (Maghsoudi *et al.* 2001, Saadoun *et al.* 2002, Auffret *et al.* 2015). The second order that had an increase after the diesel contamination, Rhizobiales, are well known for their role in fixation of nitrogen and are associated with vegetal roots that were abundant in the top soil used for that study. The 14 days of incubation was enough to see significant change in the communities structure, probably due to the growth efficiency of some species in the new condition, or the loss of growing efficiency for a large proportion of bacterial species, non-adapted to the new conditions.

Most abundant order before with the addition of diesel fuel (in relative abundance)

Actinobacteria;o__Actinomycetales	0,22050716
Alphaproteobacteria;o__Rhizobiales	0,17332596
Acidobacteria-6;o__iii1-15	0,06453044
Gammaproteobacteria;o__Xanthomonadales	0,05568127
Alphaproteobacteria;o__Rhodospirillales	0,04339379
[Spartobacteria];o__[Chthoniobacterales]	0,03198768
Sphingobacteriia;o__Sphingobacteriales	0,03159987
Planctomycetia;o__Gemmatales	0,03001068
Betaproteobacteria;o__Burkholderiales	0,02861332
Deltaproteobacteria;o__Myxococcales	0,02257576
Planctomycetia;o__Pirellulales	0,02000498
MB-A2-108;o__0319-7L14	0,01685268

[Pedosphaerae];o__[Pedosphaerales]	0,01630064
Thermoleophilia;o__Solirubrobacterales	0,0160984
Chloracidobacteria;o__	0,01417502
Alphaproteobacteria;o__Sphingomonadales	0,01382985
Gammaproteobacteria;o__Pseudomonadales	0,01198448
Thermoleophilia;o__Gaiellales	0,01086324
Acidimicrobiia;o__Acidimicrobiales	0,01023373
Solibacteres;o__Solibacterales	0,00885685
Betaproteobacteria;o__SC-I-84	0,00884062

Most abundant order without diesel fuel addition (in relative abundance)

Alphaproteobacteria;o__Rhizobiales	0,1268207
Actinobacteria;o__Actinomycetales	0,11128545
Acidobacteria-6;o__iii1-15	0,08647205
Sphingobacteriia;o__Sphingobacteriales	0,07668259
Gammaproteobacteria;o__Xanthomonadales	0,06474243
Alphaproteobacteria;o__Rhodospirillales	0,04749326
Deltaproteobacteria;o__Myxococcales	0,03967136
[Pedosphaerae];o__[Pedosphaerales]	0,03801877
[Spartobacteria];o__[Chthoniobacterales]	0,03217252
Betaproteobacteria;o__Burkholderiales	0,02614986
Planctomycetia;o__Gemmatales	0,02268799
Planctomycetia;o__Pirellulales	0,01948346
MB-A2-108;o__0319-7L14	0,01933722
Chloracidobacteria;o__	0,01790788
Thermoleophilia;o__Solirubrobacterales	0,015269
Thermoleophilia;o__Gaiellales	0,01351994
Solibacteres;o__Solibacterales	0,01285226
Deltaproteobacteria;o__Syntrophobacterales	0,01196494
Deltaproteobacteria;o__[Entotheonellales]	0,01119345
Acidimicrobiia;o__Acidimicrobiales	0,010685
Betaproteobacteria;o__	0,0105124

Figure 4: 20 most important (relative abundance) Orders with and without diesel contamination

Depth effect

Interestingly we found important changes in terms of community structure with the effect of depth. The 5 samples without diesel, taken only 6 cm under the top soils samples, were group together in a PCA, showing significant difference with the vertical dimension (depth effect) (Figure 5). The main differences were the increase of Bacteriodetes for the deeper samples and a decrease of the relative abundance of Actinobacteria (Figure 7).

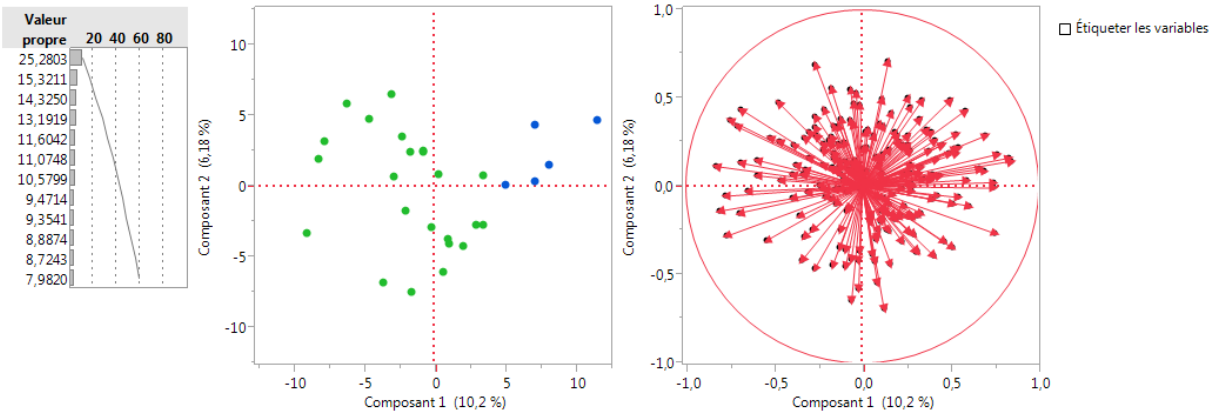


Figure 5: PCA of samples composition of uncontaminated half core. Surface samples (green dot), 6cm depth samples (blue dot).

The 6 cm deep samples taken in the contaminated half core and uncontaminated half core were more similar together than the same points on the top soil wth or without contamination. The main difference between deeper sample with or without diesel is the relative abundance of Actinobacteria. For the diesel contaminated samples, the proportion of Actinobacteria were similar then the one observed for the surface diesel contaminated sample, reducing the vertical differences in the contaminated core (Figure 6).

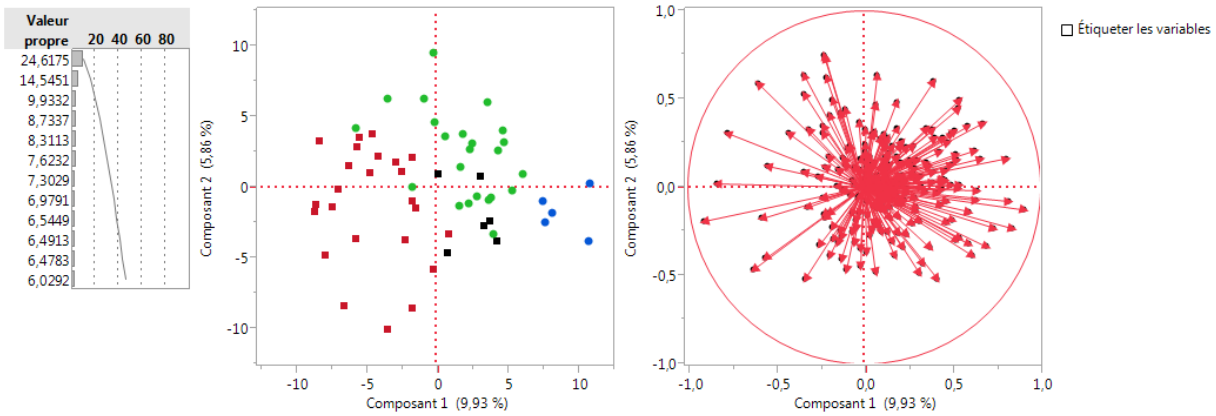


Figure 6: PCA of samples composition. Uncontaminated surface samples (green dot), 6cm depth uncontaminated sample (blue dot). Diesel surface samples (red dot). Diesel 6cm depth sample (black dot).

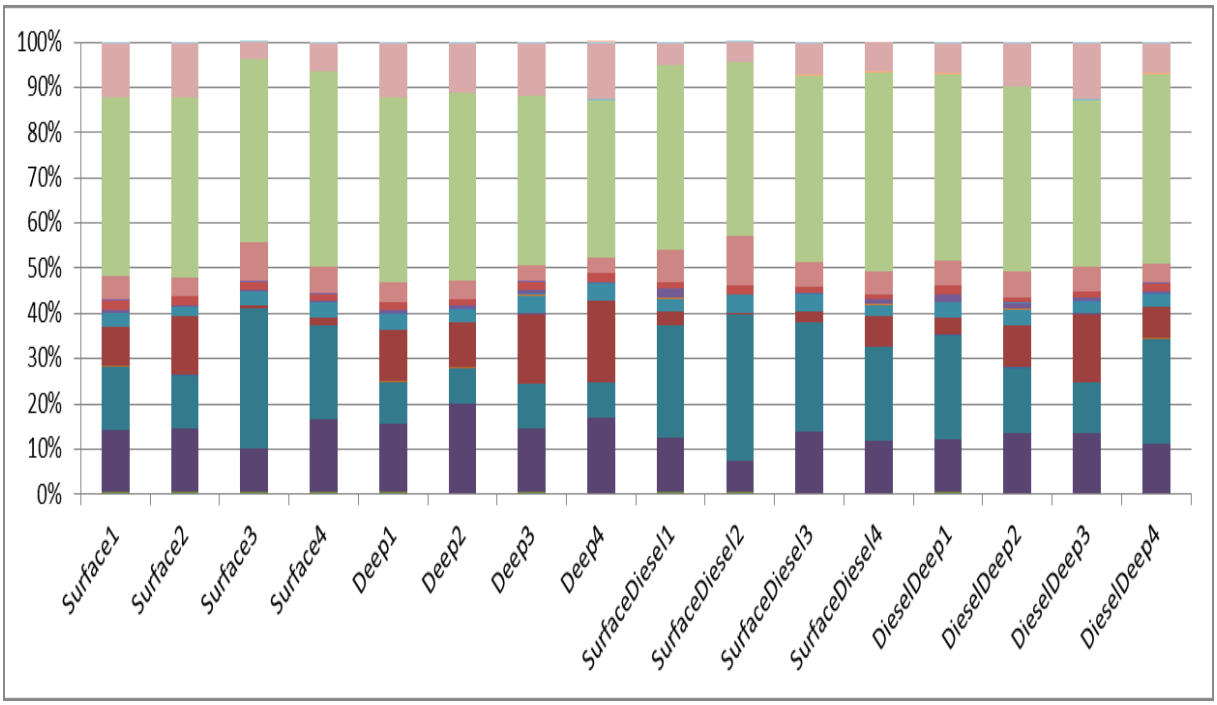


Figure 7: Phylum composition of samples taken at the same geographic position but with different depth.

Spatial analysis

Interpolation (krigging) of the spatial distribution of the main phylum for the top soil samples have shown some spatial pattern of distribution. The more abundant Phylum, Proteobacteria, show a geographical pattern different from the spatiality of the induced changes (addition of diesel), suggesting that another variable explain their spatial distribution (Figure 8d). Bacteriodetes and Acidobacteria have a spatial pattern of distribution clearly influenced by the addition of diesel (Figures 8 a, c). For the Acidobacteria we can also see in the uncontaminated half that another variable with a different spatial pattern also influence their distribution. The spatial pattern of Actinobacteria showed an increase with the diesel addition (Figure 8 b), comforting what we statistically observed and what have been described in other studies.

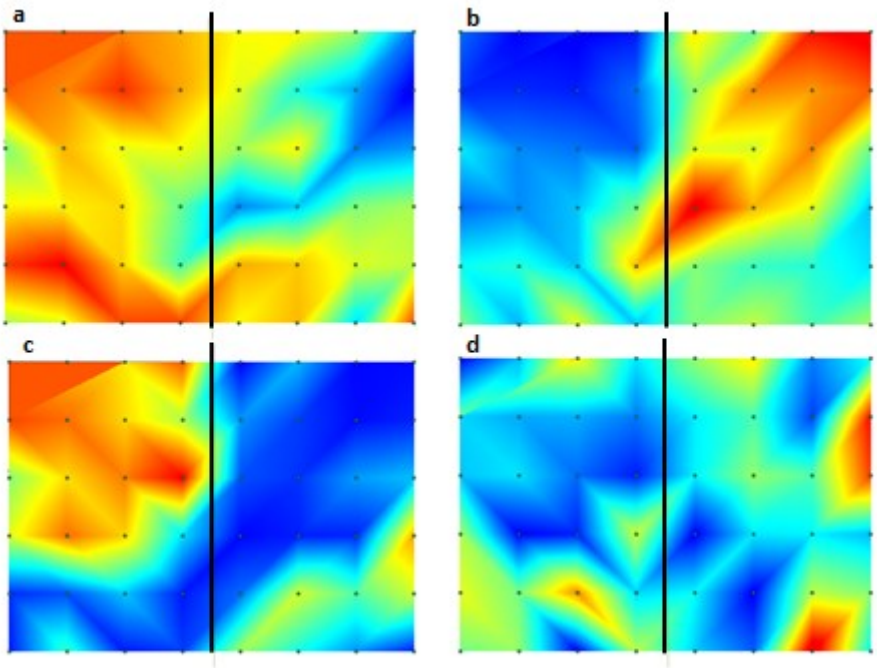


Figure 8: Interpolation of the phylum relative abundance for reconstructed core of soil, right part is without diesel and left part with diesel. Acidobacteria (**a**), Actinobacteria (**b**), Bacteriodetes (**c**) and Proteobacteria (**d**)

Distance

We observed a poor positive relationship between distance and dissimilarity (Figure 9), but the standard deviation was too large to conclude that distance is an important variable for understanding the spatial distribution of bacteria at the scale of our analysis.

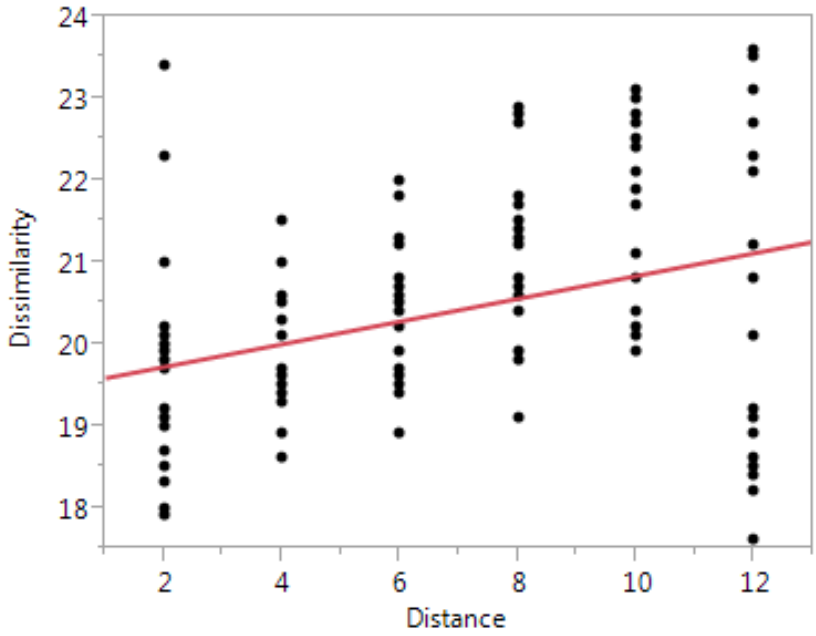


Figure 9: Dissimilarity between samples composition and distance (cm)

Discussion

The very high diversity of bacteria in soil makes it difficult to look at low taxonomic levels (Genus, Species) to understand the spatial distribution of bacteria. The high heterogeneity makes every sample different, especially when we worked with a sampling distribution at the centimeters scale or higher. Another limitation is the size of the sample used to extract DNA. In this case, we were using 250mg of soil per sample. At the scale of a bacteria, that is a significant volume that can contain up to 10^9 individual bacteria and more than 10^4 different species (Gans *et al.*, 2005; Schloss and Handelsman, 2006). Working with relative abundance at higher taxonomic levels (from Phyllum to family) reduced the number of variables to compare between samples and to group samples together. In this study, we were able to describe shared community structure at the phylum and order level for the soils with or without diesel addition. We also observed a significant difference between samples as a function of depth. That result support what has been observed in other studies (Fritze *et al.* 2000, Fierer *et al.* 2003, Stone *et al.* 2014) and highlight the importance of considering the depth in field study sampling efforts. Bacterial density and heterogeneity decrease with the depth (Eilers *et al.* 2012, Kramer *et al.* 2013) but in this study, we demonstrated that the community structure was also different at depth than at the top of the soil. The samples taken at the same depth shared more similarity than the nearer ones that were at different depths.

The important question we wanted to answer here was whether we can overcome the smaller scale spatial distribution of the variables influencing the spatial distribution of bacteria in soils by inducing a major chemical change at the scale of our sampling. We demonstrated that if the change is enough important we can see a

specific signature of the spatial distribution of bacteria at the scale of our core of soil, overcoming the micro-scale variables that drive the geographical presence of bacteria. Theoretically, we demonstrate that, in a field study, if we have a variable that have a spatial distribution at a similar spatial scale level then the sampling nest, and if the variable is a driver of the distribution of bacteria, we can overcome a part of the smaller scale effect on bacterial distribution in our system. The variable having a spatial distribution at a larger scale then the sampling are integrated in all the samples and can't be used to compare sample between them (Figure 10).

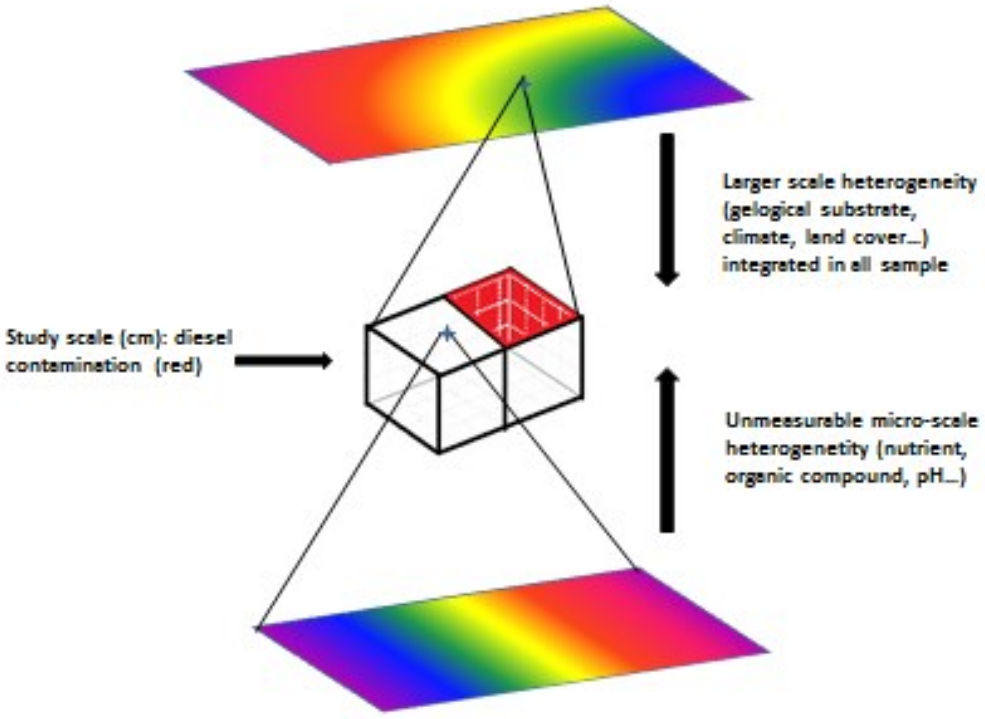


Figure 10: Multiple scale affecting soils bacterial spatial distribution

The interpolation of the relative abundance of bacterial phylum also demonstrate the feasibility of spatial analysis at larger scale then the scale of bacteria (micro-scale). Those tools and technics also allow us to interpret if the driver of the bacterial distribution happened at the same spatial scale level then the sampling. The use of GIS software (ArcGIS, GRASS, QGIS) give us the opportunity of doing spatial analysis of the different parameters to found the one corresponding with the variability of the community structure and identified the major driver at the scale of our study.

Conclusion and prerspectives

The spatial scale of bacterial habitat is way smaller than the size of a sample we use to do 16S analysis (including the step of DNA extraction). Thousands of micro habitat are include in an individual small sample and it's impossible to have the spatial information inside. To understand a part of the spatial distribution of bacteria in larger scale studies (from centimeters to continental scale) we have to identified variable that occur at the same spatial scale level and able to overcome a part of the noise coming from the smaller scale underlying variables. After 14 days of incubation, considering that a great part of death cells were still present in the extracted DNA, the differences we observed in relatives abundance, of taxa known to be present in hydrocarbs contaminated soils, were significant. Furthermore, the spatiality of variables and the sampling effort have to be consider in a 3D system as the vertical

spatial differences between community structure between sample is often higher than the horizontal variability.

List of figures

Bibliography

Alvarez, V. M., Santos, S. C., Casella, R. C., Vital, R. L., Sebastin, G. V., & Seldin, L. (2008). Bioremediation potential of a tropical soil contaminated with a mixture of crude oil and production water. *Journal of microbiology and biotechnology*, *18*(12), 1966-1974.

Auffret, M. D., Yergeau, E., Labbé, D., Fayolle-Guichard, F., & Greer, C. W. (2015). Importance of Rhodococcus strains in a bacterial consortium degrading a mixture of hydrocarbons, gasoline, and diesel oil additives revealed by metatranscriptomic analysis. *Applied microbiology and biotechnology*, *99*(5), 2419-2430.

Chikere, C. B., Okpokwasili, G. C., & Chikere, B. O. (2009). Bacterial diversity in a tropical crude oil-polluted soil undergoing bioremediation. *African Journal of Biotechnology*, *8*(11).

Curtis, T. P., Sloan, W. T., & Scannell, J. W. (2002). Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences*,*99*(16), 10494-10499.

Eilers, K. G., Debenport, S., Anderson, S., & Fierer, N. (2012). Digging deeper to find unique microbial communities: the strong effect of depth on the structure of bacterial and archaeal communities in soil. *Soil Biology and Biochemistry*, *50*, 58-65.

Fierer, N., Schimel, J. P., & Holden, P. A. (2003). Variations in microbial community composition through two soil depth profiles. *Soil Biology and Biochemistry*, *35*(1), 167-176.

Fritze, H., Pietikäinen, J., & Pennanen, T. (2000). Distribution of microbial biomass and phospholipid fatty acids in Podzol profiles under coniferous forest.*European Journal of Soil Science*, *51*(4), 565-573.

Gans, J., Wolinsky, M., & Dunbar, J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science*,*309*(5739), 1387-1390.

Janssen, P. H. (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and environmental microbiology*,*72*(3), 1719-1728.

Kelley, D. S., Baross, J. A., & Delaney, J. R. (2002). Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annual Review of Earth and Planetary Sciences*, *30*(1), 385-491.

Kramer, S., Marhan, S., Haslwimmer, H., Ruess, L., & Kandeler, E. (2013). Temporal variation in surface and subsoil abundance and function of the soil microbial community in an arable soil. *Soil Biology and Biochemistry*, *61*, 76-85.

Labbé, D., Margesin, R., Schinner, F., Whyte, L. G., & Greer, C. W. (2007). Comparative phylogenetic analysis of microbial communities in pristine and hydrocarbon-contaminated Alpine soils. *FEMS microbiology ecology*, *59*(2), 466-475.

Larose, C., Berger, S., Ferrari, C., Navarro, E., Dommergue, A., Schneider, D., & Vogel, T. M. (2010). Microbial sequences retrieved from environmental samples from

seasonal Arctic snow and meltwater from Svalbard, Norway.*Extremophiles*, *14*(2), 205-212.

Maghsoudi, S., Vossoughi, M., Kheirolomoom, A., Tanaka, E., & Katoh, S. (2001). Biodesulfurization of hydrocarbons and diesel fuels by Rhodococcus sp. strain P32C1. *Biochemical Engineering Journal*, *8*(2), 151-156.

Saadoun, I. (2002). Isolation and characterization of bacteria from crude petroleum oil contaminated soil and their potential to degrade diesel fuel.*Journal of basic microbiology*, *42*(6), 420-428.

Saul, D. J., Aislabie, J. M., Brown, C. E., Harris, L., & Foght, J. M. (2005). Hydrocarbon contamination changes the bacterial diversity of soil from around Scott Base, Antarctica. *FEMS Microbiology Ecology*, *53*(1), 141-155.

Schloss, P. D., & Handelsman, J. (2006). Toward a census of bacteria in soil.*PLoS Comput Biol*, *2*(7), e92.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., & Weber, C. F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*,*75*(23), 7537-7541.

Stone, M. M., DeForest, J. L., & Plante, A. F. (2014). Changes in extracellular enzyme activity and microbial community structure with soil depth at the Luquillo Critical Zone Observatory. *Soil Biology and Biochemistry*, *75*, 237-247.

Torsvik, V., & Øvreås, L. (2002). Microbial diversity and function in soil: from genes to ecosystems. *Current opinion in microbiology*, *5*(3), 240-245.

Tringe, S. G., Von Mering, C., Kobayashi, A., Salamov, A. A., Chen, K., Chang, H. W. Rubin, E. M. (2005). Comparative metagenomics of microbial communities. *Science*, *308*(5721), 554-557.

Chapter 3

Geographic Information System (GIS) Analyses of Large-Scale Spatial Soil Bacterial Diversity

Résumé

Les sols sont probablement, pour la fraction microbiologique, les écosystèmes les plus riches en biodiversité. Malgré de considérable efforts de séquençage d'ADN et de rARN pour de nombreux types de sol, beaucoup reste à explorer pour comprendre comment ces communautés bactériennes sont structuré, étendent leurs interactions et le rôle dans le fonctionnement des écosystèmes. La distribution spatiale des bactéries habitant le sol est hautement hétérogène, à différentes échelles, mais demeure peu connu. Des études ont toutefois démontré l'existence de liens entre la distribution spatiale des micro-organismes avec la distribution spatiale de paramètres physico-chimiques du sol (*e.g.*, relation entre le pH du sol et l'abondance relative des *Acidobacter*). Dans ce projet, nous amenons l'hypothèse que l'hétérogénéité de la composition des communautés bactériennes du sol apparaît à la même échelle que les propriétés environnementales du sol. Pour la première fois dans le cadre d'une étude terrain à grande échelle, une combinaison d'analyses par puces phylogénétique, d'analyses physico-chimiques et d'analyse spatiale à grande échelle avec des systèmes d'informations géographiques (SIG) ont été utilisée pour étudier la distribution spatiale des bactéries dans le sol, afin de comprendre la relation entre la composition bactérienne du sol et les paramètres environnementaux du sol. Les analyses multivariées des résultats des analyses phylogénétiques sur puces et des analyses physico-chimiques n'ont laissé voir aucune évidence de relations spécifiques entre les caractéristiques du sol et sa composition bactérienne, tout particulièrement pour les niveaux taxonomiques supérieurs. Par contre, avec les analyses spatiale par systèmes d'informations

géographiques, nous avons pu mettre en évidence la complexité des paramètres du sol qui contrôlent la structure des communautés microbiennes à l'échelle de très large régions avec l'exemple du pourcentage de couvert forestier versus le pH et les effets sur le phylum Acidobacteria.

Abstract

Soils are probably the most microbially diverse ecosystems on Earth. Although considerable sequencing of DNA and rRNA from different soils has been carried out and has lead to the discovery of complex microbial communities, much remains to be explored in terms of how they are structured, the extent of their interactions and their role in ecosystem functioning. The spatial distribution of bacterial communities inhabiting the soil shows high heterogeneity at different scales, but is still almost unexplored. Some studies have attempted to link the spatial distribution of soil microorganisms with soil physicochemical parameters (*e.g.*, relationship between soil pH v and *Acidobacter* abundance). In this project, we hypothesize that heterogeneity of the bacterial community composition appears at the same scale level as the heterogeneity of soil physicochemical properties. In order to understand the relationship of bacterial composition of soils (from large region in the northern France) and soil factors at different spatial scales, we applied, for the first time in a large scale study, a combination of phylogenetic microarray analysis, physical and chemical analysis and large scale geographic information system (GIS) analysis. The multivariate analysis of phylogenetic microarray results and physical and chemical analysis did not give any evidence of specific soil characteristics associated with specific bacterial community structure, especially for higher taxonomic levels. On the other hand, we were able to couple taxonomic analyses of microbial community structure and geographical information systems (GIS) to demonstrate the complexity of parameters related to shifts in community structure over large distances with the example of forest versus pH effects on Acidobacteria. This study demonstrates the power of applying multiple analytical techniques to improve our understanding of complex environments and interactions.

1. Introduction

Many drivers of the spatial distribution of plants and animals around the globe (soil nutrients, climatic conditions, altitude…) have been described (Elith *et al*., 2009, Franklin, 2010), but the drivers influencing the spatial distribution of microorganisms in soils are still poorly understood. Microorganisms play an important role in soil ecosystems by contributing to biogenic transformation of carbon, nitrogen and phosphorus (Paul and Clark 1989). Several molecular biology methods (including DNA-DNA re-association, DNA cloning and sequencing and fingerprinting approaches) have been used to estimate bacterial diversity in soil (Frostegard et al 1993, Kowalchuk et al 1998, Muyzer et al 1993, Torsvik et al 1990). These estimates are generally based on one or a small number of soil samples with neither spatial nor soil characteristic variations. Yet, the heterogeneity of abiotic and biotic characteristics in soils has been observed at the micro to macro-scale (Beare et al, 1995, Ramette et al. 2007). At the micro-scale, some studies have highlighted the role of micro-niches (Grundmann et al. 2004), aggregates (Vos et al. 2013), soil structure (Remenant et al. 2009) and organic matter (McCabe et al. 2011) on the variable microbial density in soil. At the meso-scale (field scale), principally vegetation cover (Uroz et al. 2010) and nutrient abundance (Philipot et al. 2009, Wessen et al. 2011) have been shown to correlate with phylogenetic and functional microbial diversity in soil. At the macro-scale, not many significant correlations with abiotic characteristics have been elucidated other than the correlation of the relative abundance of *Acidobacteria* with differences in soil pH (Fierer et al. 2006, Griffiths et al. 2009). While these studies demonstrate correlations, they do not necessarily show causation. In the case of pH variations, soil organic matter type also correlates with low pH forest soils (and their associated resins) at one extreme and high pH

grasslands at the other. Thus, the proportion of forest soil in the drainage area of the sampled point might be the critical driver of *Acidobacteria* abundance. Part of the difficulty with phylogenetic analyses based on the sequencing of relatively variable regions is the lack of more conserved regions within the analyses for uncultivated bacteria (Stackebrandt and Ebers, 2006; Mizrahi-Man et al., 2013). One approach is to sequence several regions of the 16S rRNA gene and another is the use of tiered phylogenetic microarrays with probes from different parts of the 16S rRNA gene (Claesson et al., 2009). Phylogenetic microarrays were developed to identify bacterial species and to assess bacterial diversity (Cho and Tiedje 2002, Sanguin et al 2006). A single array can contain several thousand different 16S rRNA gene sequences, and thus identify different taxa within bacterial communities simultaneously. Microarrays can also contain probes that target uncultivated bacteria (currently the largest proportion of soil bacterial communities) at different taxonomic levels. In this study, we combined the use of a high density microarray and geographic information system (GIS) to study the spatial distribution of bacterial diversity in soil at a regional scale. Using this combined approach, we examined the relationship between bacterial community diversity and the physicochemical properties and macro-biotic characteristics that drive soil chemistry.

2. Materials and methods

2.1. Soil samples and soil characteristics

The different soil samples (64) used in this study consisting of composite samples from within a square of 256x256 Km in the middle of France (Figure 1) and were part of a larger set of samples covering most of France (Dequiedt et al., 2009). This square was divided into 64 smaller squares of 32x32 Kmfrom which 25 subsamples were taken from the center of the square at a depth of 10 cm of depth. These subsamples were pooled to form a composite sample in order to have a more stable and representative bacterial community and physical-chemical characteristics. . All composite samples were dried at 25°C, homogenized and sieved at 20µm. These soils were characterized by a range of organic matter concentrations (carbon, nitrogen, and available phosphorus (supplemental data in annex files)), physical (sand and clay percent) and chemical (pH) characteristics, and soil uses.

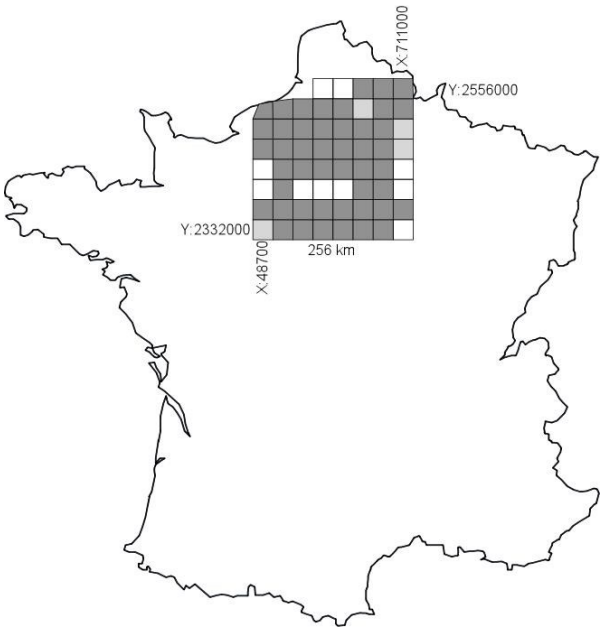


Figure 1: Geographical position of soils sampled in the center of France. X and Y are GPS positions of sub-squares flanking the square. White squares were not sampled and light gray squares were not included in the analysis.

2.2. DNA extraction and amplification of 16S rRNA gene (*rrs*)

Total DNA extraction and purification were carried out according to the protocol described by Ranjard et al. (Ranjard et al 2001). All DNA was prepared with a final concentration of 6.67ng/µl. PCR amplification was performed using pA (forward) and pH (reverse) primers (Edwards et al 1989). In order to transcribe the PCR products to 16S rRNA antisense strand; the reverse primer was modified by incorporating a T7 promoter during production. Amplifications were carried out using the Hot Start Mix RTG kit (GE Healthcare, UK Limited) for a total mixture of 25µl. For each reaction, 1.25 µl (0.5mM final) of each primer, one batch of Hot Start Mix RTG and 2 µl of DNA were used. For all amplifications, the following cycle was used: 2min at 94°C, 35 cycles composed of 30s at 94°c, 30s at 58°c and 45s at 72°c. The amplification finished with 5 min at 72°c.

2.3. Labeled of target 16 rRNA genes

PCR products were then purified with the GFX PCR DNA and Gel Band Purification kit (GE Healthcare, UK Limited). *In vitro* transcription was carried out at 37°C during 4 hours in 20 µL reactions that contained 8 µL of the purified PCR product (50 ng.µL^{-1}) and 12 µL of the following mix: T7 RNA buffer (5X), DDT (100 mM), 10 mM of each of the four NTPs, RNasin (40 U.µL^{-1}), T7 RNA polymerase (1 µL) and UTP-Cy3 (5 mM). During transcription, Cy3-UTP (a fluorescent dye that emits light at 532nm) is incorporated to label RNA.

2.4. Microarray preparation

RNA was purified using the Quiagene RNeasy mini Kit according to the manufacturer's instructions and quantified with a nanophotometer before undergoing chemical fragmentation by addition of 5.7 µL of a Tris Cl (1mM) and $ZnSO_4$ (100mM) mix. Samples were incubated for 30 minutes at 60°C and fragmentation was stopped

by placing the tubes on ice. EDTA (500mM) was added to each tube (1.2 µL) followed by 1 µl RNAsin (40 $U.\mu L^{-1}$) after a minute incubation period at 25°C. The RNA solution was then diluted to 5 $ng.\mu L^{-1}$ and a hybridization mix was prepared (v/v ratio) in a 50 µL reaction with 2x GeX Hyb Buffer (Agilent). A total of 100 ng of RNA were then placed on the slide and incubated at 60°C for 4 h in the Agilent Hybridization Oven.At the end of hybridization, microarrays were removed from the hybridization chambers and washed according to the manufacturer's instructions.

2.5. Probes of 16 rRNA microarray and Microarray design

The Agilent Sureprint Technologies microarray format was used, and consisted in 8 identical blocks of 15,000 spots each on a standard glass slide format 1" x 3" (25mm x 75mm). Each spot is formed by *in situ* synthesis of 20-mer oligoprobes that occur at least in triplicate within each block. The syntheses were completed by Agilent probes composed of negative and positive control and probes for locating Agilent Gal file. Probes were designed to target the *rrs* gene at different taxonomic levels (1469 genera, 286 families, 118 orders, 57 classes, 36 Phyla based on NCBI taxonomy) from the *Bacteria* and *Archaea* phylogenetic tree using the ARB software package (phylogenetic microarray target) (for probes see http://www.genomenviron.org/Research/Microarrays.html) . *Proteobacteria* were the most represented group, with alpha-*Proteobacteria* representing 29.6% of all probes. The nine most represented bacterial groups on the microarray were respectively: *Bacilli*, *gamma-Proteobacteria*, *Flavobacteria*, *Actinobacteria*, *beta-Proteobacteria*, *Clostridia*, *Sphingobacteria* and *Bacteroidetes*. In order to be sensitive to uncultured soil bacteria, the microarray contained probes (5.7% of total) targeting uncultured bacteria at the genus (or species) level. The microarray also targeted (5.9% of probes) taxa referenced in the data base from environmental samples.

2.6. Signal and data capture

Microarrays were scanned using an Axon GenPix 4100b scanner (Molecular Devices, USA) at 5 µm resolution and at a PMT gain of 500. Spots with aberrant or saturated signals were individually removed from the analyses. The files were exported as GPR files. Total intensity of each array were extracted from GPR files and used for further analyses.

2.7. Microarray normalization

In order to normalize data between microarrays, we use the linear method based on adjustment of quintiles of total intensity between blocks from the same microarray and between microarrays. The algorithms used are described in the LIMMA package implemented in R (Xia et al., 2005).

After the microarray data was normalized, two criteria were determined to select the spots corresponding to a positive signal. The first one was the threshold value for positive hybridization (PH), based on the values of signal to noise ratio (SNR) above 3. The second criterion for selecting a positive signal was when the value of total intensities was higher than the Agilent negative control ((-) 3xSv1). We considered taxa as present in samples when targeted probes fulfilled these two criteria. For the probes targeting microorganisms at the level of phylum, class or family, the same criteria were also used.

2.8. Statistical analysis

Data analysis was carried out using statistical packages implemented in R. Microarray data were normalized with the quantile method described in the R package LIMMA,. Principal Component Analysis (PCA) and Canonical Correspondence Analysis (CCA) were performed using the CANOCO software

version 5. For the Mandel test, the Jacquard and Euclidian distances were calculated with the Vegan package implemented in R-project.

2.9. GIS analysis

The first step was to determine the drainage area for each geographical point. Twenty-five samples were taken for each point and all the soil areas that contribute to the organic matter (and other nutrient) load for each of the sampled points were determined. We analyzed digital elevation model (DEM) imagery with a resolution of 30 m.pixel^{-1} (GRASS 6.4) to estimate specific drainage areas. Several SRTM (High-resolution topographic data generated from NASA's Shuttle Radar Topography Mission, SRTM) images were patched together to create a single map of the entire region. Using the GRASS tool, we calculated the drainage area for each sampled point. The resulting raster maps were hand corrected using NVIZ 3D visualization tool (GRASS) and converted into vector maps (Figure 2).

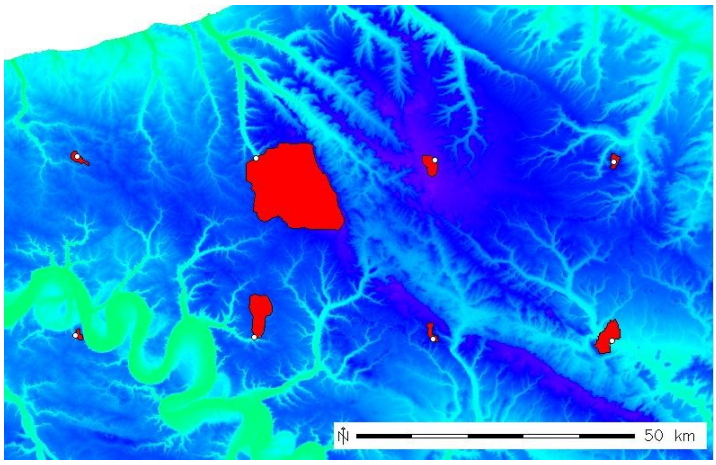


Figure 2: Drainage area polygons derived from DEM

The different types of land cover were assessed using nine matrices of Landsat 7 satellite images. The snapshots were interpolated to cover the whole region using

GRASS 6.4 and then analyzed for vegetation cover. Ten classes of reflectance were created, each one representing a particular land cover (forested area, shrub land, herbaceaous, uncultivated agricultural soil, cultivated agricultural soil, lawn or urban green, concrete or asphalt, bare soil, peat land and water). The resulting regional image of the land cover classes was then transformed into a vectorial projection (ArcMap) and was augmented with a layer representing the drainage area contours for the calculation of specific cover areas (Beaulne, 2008).

3. Results and discussion

3.1. Microarray

In order to characterize the bacterial communities within the different squares sampled, a high density microarray was used. Of the 3195 probes targeting microbial taxa on the microarray, a total of 1693 probes showed a positive hybridization with at least one sampled soil, representing more than half of the designed probes. Among them, 18% of the probes gave a positive signal with DNA from all the squares, indicated that the sampled soils shared a core microbial community. *Alpha-Proteobacteria* was the most represented taxon with 7.8% of positive probes. The other most dominant taxa based on probe hybridization were *Bacteroidetes*, *Firmicutes*, *gamma-Proteobacteria* and *Actinomycetes*.

All soil samples covered at least 47% of the total positive fraction on the microarray, except soil S.624 with 33% of positive probes.. Soil S.749 and S.626 had the highest hybridization richness (HR) with 80% and 76% of positive probes, respectively, suggesting that the microarray is representative of the bacterial diversity of the sampled soils.

3.2. Microarray Validation

The microarray was designed to include both a high number of probes and probes targeting specific taxa including uncultivated bacteria and rare soil taxons such as pathogens, marine bacteria or others. The specificity and sensitivity of the microarray was validated by focusing on the detection of these probes in our soil samples. About 10% (167 probes) of the probes showed only one positive hybridization with the 16 rDNA gene from the 47 soils. All of these probes targeted microbial populations at the genus level (except five probes: one targeting *alpha-Proteobacteria* at the family level (a3), two targeting *Archeobacteria* at the phylum level (Arch1), one targeting *Bacteroidetes* class level (Bcd2)). Four probes (PrbEC3132, PrbEC0527, PrbEC2626 and PrbEC2517) targeting *Bacteroidetes* species showed a positive signal for only four soil samples (S.392, S.743, S.749 and S626). All the genera detected by these probes are represented by a single species recently isolated from freshwater that might consequently be absent or found at very low frequencies in soils. The proportion of positive probes increased by 10% when the threshold was fixed at five positive hybridizations, confirming that many taxa were detected at very low frequencies in the 47 soils. On the other hand, there was a significant positive correlation between the number of all positive probes and positive fraction >5. This finding indicated that the detection of these infrequent species could be density dependent and that the microarray used was highly sensitive.

3.3. Relationship between physical and chemical characteristics of soils

In order to analyze the relationship of physical and chemical characteristics between soils, a multivariate analysis was performed. The first component (36.7% of total inertia) separated soils rich in clay and fine silt with high alkalinity and cation

exchange capacity (CEC), rich in metals (Fe, Al, Mn,) and cations (K, Mg, Ca) from soils with high sand (coarse and fine) and C/N ratio. The second component, which explained 18.4% of the total variance, separated soils with high carbon and nitrogen content from soils rich in silt and potassium. Soils did not cluster with their neighboring samples, therefore the distance between squares was sufficient to ensure that the physical and chemical characteristics did not correlate between geographically grouped samples. The separation between agricultural, grassland and forested soils was not significant based on the physico-chemistry of the soils (Figure 3).

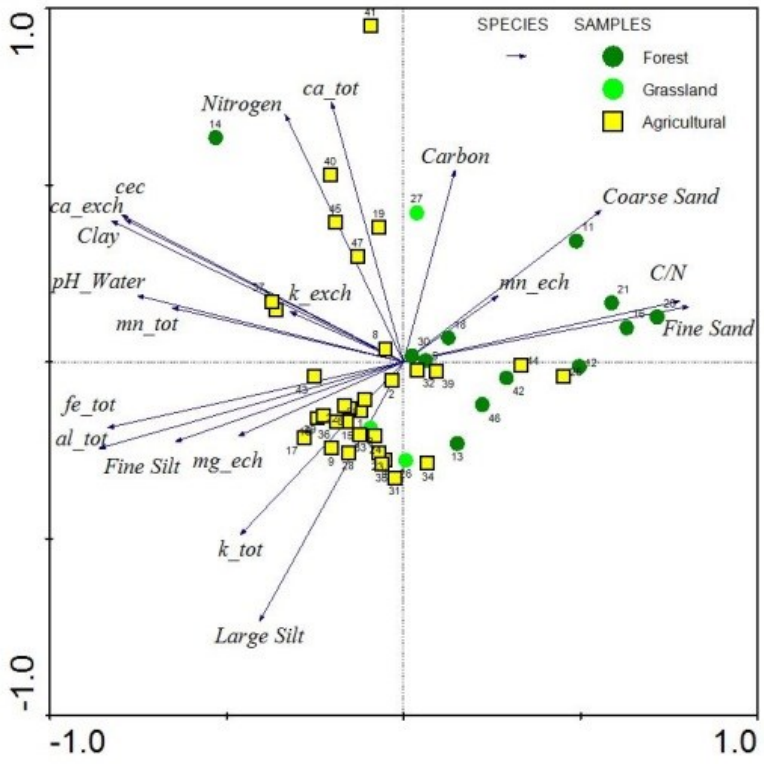


Figure 3: Principal component analysis of the physicochemical soil parameters

3.4. **Bacterial community structure in relation to soil characteristics**

In order to characterize the soil bacterial communities, PCA of the normalized microarray data were carried out at different taxonomic levels. The phylum analysis did not show a separation based on land cover type (Figure 4). The PCA at the genus level demonstrated a slight separation between most of the agricultural soil samples and the other samples (Figure 5). The first component (77.1% of relative variance) separated soils without clustering samples by the type of soils. The second component (5.3% of relative variance) differentiated most (eight of total of twelve) of the forest soils from other soils. These PCA analyses (Figure 4 and 5) showed a different pattern from those carried out with physical and chemical characteristics (figure 3). The global distribution of the relative intensities of the 1693 probes did not demonstrate a particular pattern. In contrast, those targeting certain bacterial populations had a pattern relatively similar to the distribution of physical and chemical characteristics of soil. For instance, probes targeting *beta-Proteobacteria* and *Acidobacteria* were distributed following the scales of the first and second components, respectively, indicating that these taxonomic groups could be strongly structured by soil characteristics. We found better correlations with land use at lower taxonomic levels, as was recently reported by Philippot et al. (2010).It is likely that the diversity at higher taxonomic levels is too high to cluster soils.

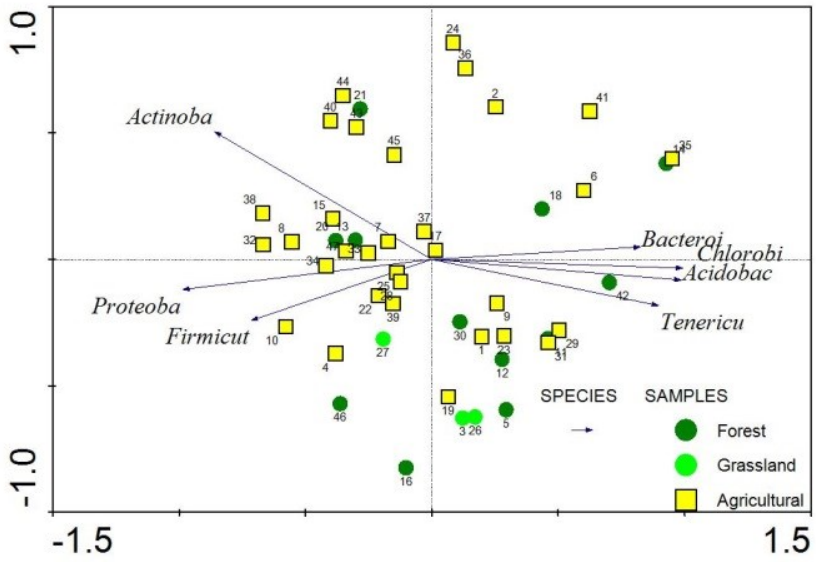


Figure 4: Principal component analysis of the microbial taxonomy at the phylum level based on phylogenetic microarray results

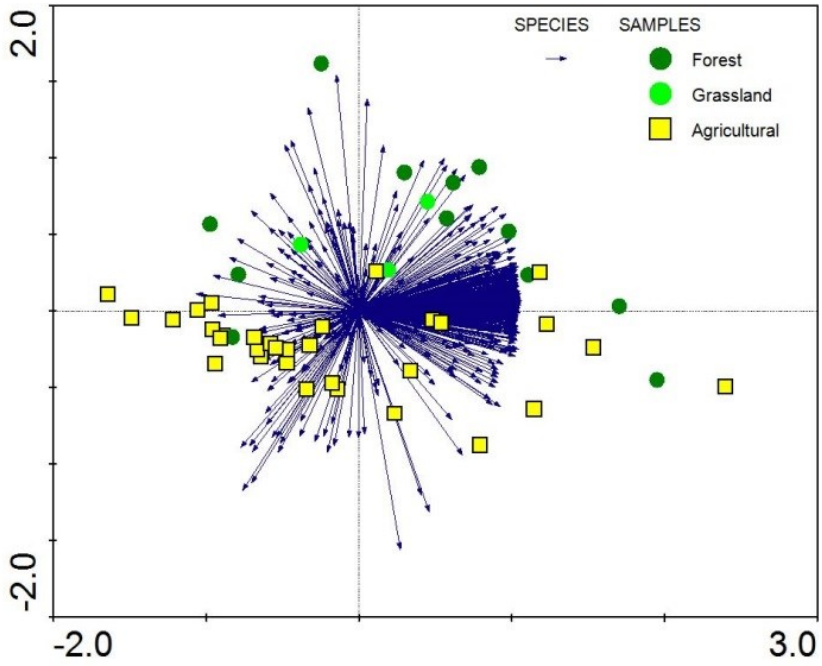


Figure 5: Principal component analysis of the microbial taxonomy at the genus level based on phylogenetic microarray results

3.4.2 **Relationships between soil microbial community structure and soil physico-chemistry**

In order to understand the relationship between soil bacterial populations and physical and chemical characteristics, Canonical Correspondence Analysis (CANOCO) was performed between the distance matrices of microarray data and soil factors. At the phylum, class and order levels (figure 6), we did not observe any significant correlation between the bacterial composition of the samples and the physico-chemical parameters of the soils. Only some probes at the class level were significantly correlated with soil physico-chemical parameters (*Actinomycetes*, *Firmicules* (*Bacilli*), *Beta-Proteobacteria*, *Acidobacteria*). For the lowest levels, family and genus (figure 7 and 8), we obtained highly significative correlations between bacterial community composition and physico-chemical parameters (p-value = 0.0002). Only one of the forested soils, dot 14 (S.198) is not grouped with the others, suggesting that the bacterial community in forested soils is mainly driven by low pH, high carbon, sand fraction and C/N.

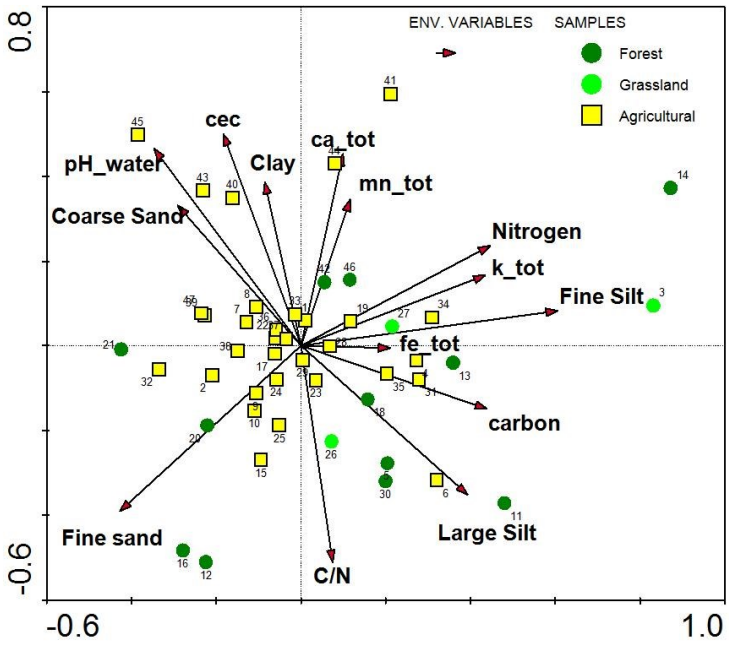


Figure 6: Canonical correspondence analysis at phylum level composition and soil physicochemical parameters (p-value = 0.2)

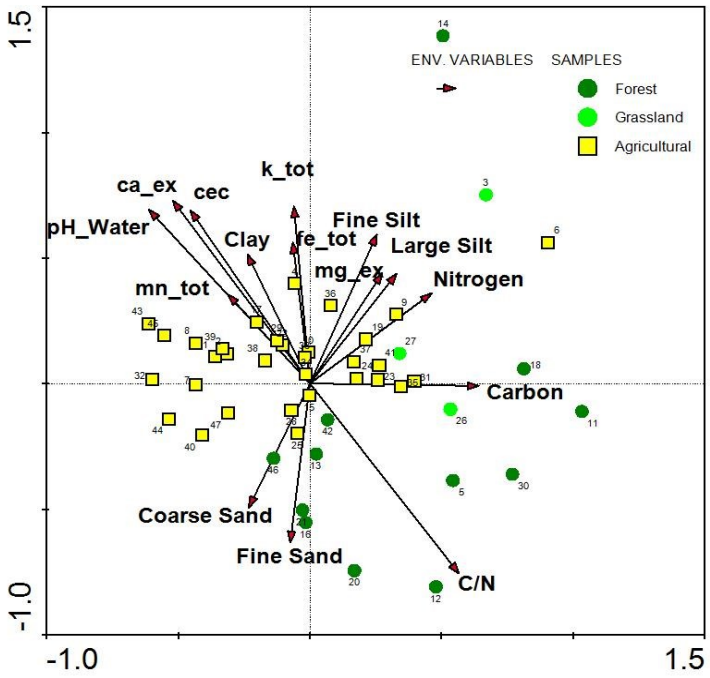


Figure 7 : Canonical correspondence analysis famiy level composition and soil physicochemical parameters (p-value = 0.0002)

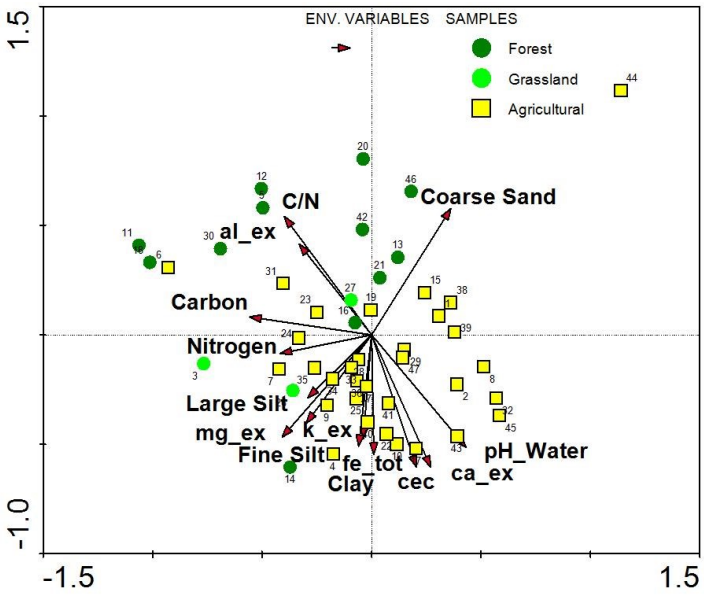


Figure 8: Canonical correspondence analysis at genus level composition and soil physicochemical parameters (p-value = 0.0002)

3.5. GIS analysis and large scale drivers of the spatial distribution of soil bacteria

Although pH has been identified as one of the variables with the most influence in driving microbial community structure in soils at large spatial scales, it is difficult to model given that the spatial variability can be important at smaller scales. It may therefore be useful to identify other proxies for pH that can be more easily integrated into large-scale models. In order to test whether GIS analysis can be used to predict the spatial distribution of microbial community structure, we used variable wood land as a proxy for pH, since pH can be related to forested soils. We then applied GIS analysis to model the abundance of *Acidobacteria*. In the previous section, the forest soils were only identified by the land cover directly over the sampled points. To integrate temporal change and matter mobility, we chose to work on the drainage area of each of the sampled points, which was calculated by merging the woodland classes representing lignin matter from the landsat 7 reflectance images and observed a relation with the density of *Acidobacteria* probes (figure 9). In figure 9, the size of each circle represents the number of Acidobacteria hits measured by the taxonomic microarray. The green circles represent samples where either more than 30% of the drainage area is covered by wood land or where the sample was sampled directly from a forest soil. The brown circles represent samples collected from sites with less than 30% wood land or not sampled from a forest soil.

Using simple environmental variables derived from available geographic data and satellite imagery, we were able to link the land cover and *Acidobacteria* presence. The advantage of this approach is that it can be applied to a wide regional scale and

is functional even for a large scale study of environments with high heterogeneity and complex environmental parameters.

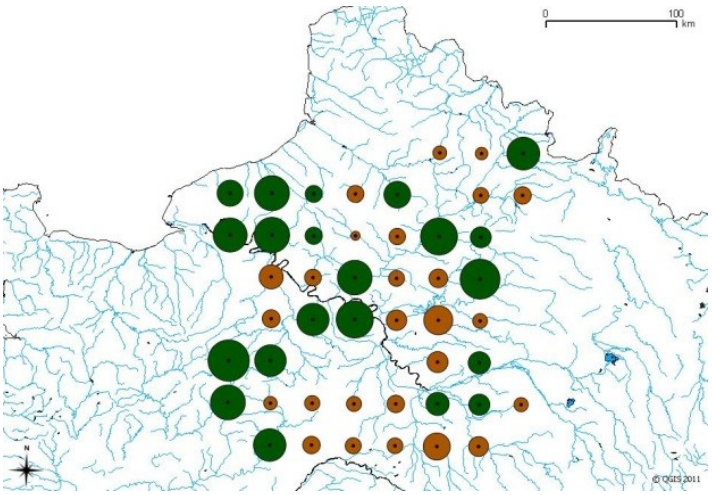


Figure 9: Relationship between *Acidobacteria* (proportional to circle size) and forest soils (green circle). Brown circles are grasslands or cultivated soils.

4 Discussion

As both the diversity of bacteria (up to 10^7 species per gram of soil) and density are relatively high (up to 10^9 per gram of soil) in soils, the understanding of the spatial distribution of bacteria in soils is critical to understand their part on the global biogeochemical cycles. Some studies have attempted to link environmental parameters with soils microbial communities, but only pH has been highlighted a strong driver of soil bacterial communities, principally by controlling the density of the *Acidobacteria* class (Fierer et al. 2006, Griffiths et al. 2009). We chose to focus on

the *Acidobacteria* to compare the result obtained by the phylogenetic microarrays analysis and the GIS analysis performed over the entire region. Working at the scale of an entire region, we have to evaluate parameters that will have an effect on the large scale heterogeneity of soils pH. The spatial heterogeneity of the chosen variable have to be at the same large spatial scale then the one of the soil sampling. The variable “forested soils” was the more scale-related for the purpose of our study. While forested soils are well known to be positively correlated with the soil pH (Binkley *et al.*1989), the transport of forest soil organic matter depends on the drainage. We chose to look at the drainage area of each sampled point to determine if it was forested or not. The result we obtain was similar to a previous study comparing pH and *Acidobacteria* relative abundance from 16S rRNA gene sequencing (Fierer et al. 2006, Griffiths et al. 2009). That allowed us to go further and model the distribution of the *Acidobacteria* on large scale region based on landscape analysis of the land cover and hydrological modelling. The forest effect was greater than the soil pH effect, that was not visible in our statistical analysis at class level. By increasing our knowledge on the drivers of the distribution of bacteria in soils, we will be able to model the distribution of bacteria in soils and to model the changes that could appear in the communities structure after environmental changes or perturbations using a coupled microbial community/GIS approach.

The physico-chemical parameters analyzed for the samples have shown that they are certainly related to the microbial community structure, but due to the diversity inside each class level, we are not able to identify specific soils characteristic influencing the spatial distribution of the bacteria at the class level. We identified pH as a strong driver for the global community structure, but mostly with Acidobacteria

and *Actinobacteria*. To analyze the spatial distribution at a lower level was not possible due to the high diversity and the relatively low number of samples (47).

5. Conclusion

This study represent the first successful experiment to link phylogenetic microarrays, physical and chemical analysis, multivariate analysis tools and large-scale GIS analysis in order to determine drivers of the spatial distribution of bacteria in soils. We were able to couple taxonomic analyses of microbial community structure and geographical information systems (GIS) to demonstrate the complexity of parameters related to shifts in community structure over large distances with the example of forest versus pH effects on Acidobacteria. This approach with the data rich satellite images can be applied to discover links between bacterial community structure and environmental parameters and soil use. By continually refining our knowledge of the drivers of the spatial distribution of bacteria, we will be able to model the distribution of the bacteria at different spatial scales and understand their variations as a function of the soil environmental parameters and use.

Bibliography

Beare MH, Coleman DC, Crossley Jr DA, Hendrix PF, Odum EP (1995). A hierarchical approach to evaluating the significance of soil biodiversity to biogeochemical cycling. *Plant and Soil*, 170: 5-22.

Beaulne, J. S. (2008). Modélisation de la présence de mercure dans la chair des brochets et des dorés des lacs les plus pêchés de l'Abitibi-Témiscamingue: une approche par les systèmes d'information géographique. Master Thesis Report, UQAM

Binkley, D., C.T. Driscoll, H.L. Allen, P. Shoeneberger, and D. McAvoy. (1989). Acidic Deposition and Forest Soils: Context and Case Studies in the Southwestern United State. Manual. Edwards Brothers Publication, Michigan.

Cho JC, Tiedje JM (2002). Quantitative detection of microbial genes by using DNA microarrays. *Appl Environ Microbiol.* 68: 1425-1430.

Dequiedt, S., Thioulouse, J., Jolivet, C., Saby, N., Lelievre, M., Maron, P. A., Ranjard, L. (2009). Biogeographical patterns of soil bacterial communities. *Environmental microbiology reports*, *1*(4), 251-255.

Edwards U, Rogall T, Blocker H, Emde M, Bottger EC (1989). Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res*. 17: 7843-7853.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677.

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 626-631.

Franklin, J. (2010). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Frostegard A, Tunlid A, Baath E (1993). Phospholipid Fatty Acid Composition, Biomass, and Activity of Microbial Communities from Two Soil Types Experimentally Exposed to Different Heavy Metals. *Appl Environ Microbiol*. 59: 3605-3617.

Griffiths G., Thomson R.I., James B.C., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of British soils. *Environmental microbiology*, *13*(6), 1642-1654.

Grundmann, G. L. (2004). Spatial scales of soil bacterial diversity–the size of a clone. *FEMS Microbiology Ecology*, *48*(2), 119-127.

Kowalchuk, G. A., Bodelier, P. L., Heilig, G. H. J., Stephen, J. R., & Laanbroek, H. J. (1998). Community analysis of ammonia-oxidising bacteria, in relation to oxygen

availability in soils and root-oxygenated sediments, using PCR, DGGE and oligonucleotide probe hybridisation. *FEMS Microbiology Ecology*, *27*(4), 339-350.

Lavelle P (1997). Faunal activities and soil processes : adaptative strategies that determine ecosystem function. *Advances in ecological recearch*. 27: 93-130.

McCabe, D. P. (2011). Spatial location of active soil bacteria and their association with soil organic matter fractions (Doctoral dissertation, University of Reading).

Muyzer G, Waal ECd, Uitterlinden AG (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol*. 59: 695 -700.

Ovreas L, Torsvik VV (1998). Microbial Diversity and Community Structure in Two Different Agricultural Soil Communities. *Microb Ecol*. 36: 303-315.

Paul EA, Clark FE (1989). Soil microbiology and biochemistry: San Diego.

Philippot, L., Andersson, S. G., Battin, T. J., Prosser, J. I., Schimel, J. P., Whitman, W. B., & Hallin, S. (2010). The ecological coherence of high bacterial taxonomic ranks. *Nature Reviews Microbiology*, *8*(7), 523-529.

Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiology Ecology*, *62*(2), 142-160.

Ranjard L, Poly F, Lata JC, Mougel C, Thioulouse J, Nazaret S (2001). Chracterization of bacterial and Fungal Soil Communities by Automated Ribosomal Intergenic Spacer Analysis Finderprintints : Biological and methodological Variability. *Appl Environ Microbiol*. 67: 4479-4487.

Remenant, B., Grundmann, G. L., & Jocteur-Monrozier, L. (2009). From the micro-scale to the habitat: assessment of soil bacterial community structure as shown by soil structure directed sampling. *Soil Biology and Biochemistry*,*41*(1), 29-36.

Sanguin H, Remenant B, Dechesne A, Thioulouse J, Vogel TM, Nesme X et al (2006). Potential of a 16S rRNA-based taxonomic microarray for analyzing the rhizosphere effects of maize on Agrobacterium spp. and bacterial communities. *Appl Environ Microbiol*. 72: 4302-4312.

Smalla K, Wieland G, Buchner A, Zock A, Parzy J, Kaiser S et al (2001). Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Appl Environ Microbiol*. 67: 4742-4751.

Thioulouse J, Chesse D, Dolédec S, Olivier JM (1997). ADE-4: a multivariate analysis and graphical display software. *Stat Comput*. 7: 75-83.

Torsvik V, Goksoyr J, Daae FL (1990). High diversity in DNA of soil bacteria. *AEM.* 56: 782-787.

Uroz, S., Buée, M., Murat, C., Frey-Klett, P., & Martin, F. (2010). Pyrosequencing reveals a contrasted bacterial diversity between oak rhizosphere and surrounding soil. *Environmental Microbiology Reports*, *2*(2), 281-288.

Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of bacterial diversity in soil. *FEMS microbiology reviews*, *37*(6), 936-954.

Wessén, E., Söderström, M., Stenberg, M., Bru, D., Hellman, M., Welsh, A. & Hallin, S. (2011). Spatial distribution of ammonia-oxidizing bacteria and archaea across a 44-hectare farm related to ecosystem functioning. *The ISME journal*, *5*(7), 1213-1225.

Xia X, McClelland M, Wang Y (2005). WebArray: an online platform for microarray data analysis. *BMC Bioinformatics.* 6: 306.

Yergeau E, Kang S, He Z, Zhou J, Kowalchuk GA (2007). Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. *Isme Journal.* 1: 163-179.

List of figures

Chapter 4

Spatial and Temporal Analysis of Bacterial Diversity in Lake Sediment

Résumé

Les sédiments sont, avec les sols, les écosystèmes les plus diverses sur la planète. La distribution spatiale des communautés bactériennes habitant les sédiments est hautement hétérogène à différentes échelles spatiale et cette variabilité spatiale a été très peu explorée. Des études ont démontré des liens entre la distribution spatiale des bactéries et différents paramètres physico-chimique du sol (*e.g.* relation entre le pH du sol et l'abondance relatives des *Acidobacter)*. Dans ce chapitre nous amenons l'hypothèse que l'hétérogénéité spatial des communautés bactériennes varie à la même échelle que l'hétérogénéité spatiale des propriétés chimiques des sédiments. Nous nous intéresserons à la diversité bactérienne des sédiments à macro-échelle (Km). Selon la littérature, les variables physico-chimiques qui peuvent avoir une incidence sur la distribution spatiale des bactéries à cette échelle sont la couverture du sol des bassin-versants, le climat, le pH et la salinité. Pour tester cette hypothèse nous avons examiné la distribution spatiale des bactéries dans les sédiments du Lac Chilika (Inde) le deuxième plus grand lac d'eaux saumâtres au monde, le plus grand d'Asie. Soixante-douze échantillons (24 stations, 3 saisons – Hiver, pluie et été) de sédiment du lac Chilika furent analyser par pyroséquensage 16S rRNA. L'analyse de la couverture de surface a été réalisé avec des images satellites (Landsat) et des modèles d'élévation digitale à l'aide des logiciels GRASS et QuantumGIS. Un large spectre d'analyse physico-chimique (e.g. pH, turbidity, salinity, conductivity, nitrate) furent réalisé sur l'eau et les sédiments pour chaque station d'échantillonnage et pour chaque saison. Apres un ouragan qui dévasta la région en 2013, des échantillons supplémentaires furent collectés afin de mesurer l'influence de phénomènes climatiques extrêmes (tempête tropicale) sur la distribution spatiale des bactéries dans les sédiments. Les résultats des analyses de

l'rRNA 16S et des paramètres physico-chimiques interpolés géographiquement démontre clairement une relation spatiale entre la distribution de paramètres physico-chimiques (salinité), géomorphologiques (drainage, fermes aquatiques) et la distribution des communautés microbiennes habitant le sédiment.

Abstract

Sediment and soils are among the most microbial diverse ecosystems on the Earth. The spatial distribution of bacterial communities inhabiting the sediments is highly heterogeneous at different spatial scales, but is still mostly unexplored. Several studies suggest links between the spatial diversity of soil microorganisms and soil physicochemical parameters (*e.g.*, relationship between soil pH and *Acidobacter* abundance). In this project, we postulate that heterogeneity of the bacterial community composition varies at the same scale of the heterogeneity of sediment chemical properties. Here, we focused on the large spatial scale (km) diversity in a brackish water lagoon. The large scale physical and chemical characteristics that we hypothesize influence microbial communities in lake sediment at the kilometer scale are land cover, climate, pH, and salinity. We tested this by examining the spatial and temporal distribution of bacteria and physical and chemical parameters in sediment of the second largest brackish lake in the world (Chilika Lake, India). Seventy-two samples (24 stations, 3 seasons winter, rainy and summer) of sediments from Chilika Lake were analyzed by 16S rRNA gene pyrosequencing. Land cover analyses were performed using satellite images and a digital elevation model with geographic information system (GIS), and a large set of physico-chemical analyses (e.g., pH, turbidity, salinity, conductivity) were also performed on the water column over the sediment. After a very severe cyclonic storm (Phailin) passed near the lagoon in 2013, more samples were collected to see the impact of the tropical storm on the spatial and temporal distribution of bacteria in the sediment. The results of 16S rRNA gene analysis and physical and chemical parameters used with the spatial analysis demonstrated clear spatial relationships between physico-chemical parameters

(salinity), land surfaces (drainage area) and the distribution of sediment microbial communities.

Introduction

Microorganisms have major role in global biogeochemical cycles of carbon, nitrogen, and phosphorus. With the cellular production rate of all microorganisms on earth estimated at over 10^{30} cells/yr. (Withman *et al.,* 1998), the turnover rates for geochemical cycling could inarguably be microbially driven. The largest environmental reservoirs of microorganisms in decreasing order are soils and the marine sediments. In addition to the high number of microorganisms, soils also have relatively higher heterogeneity of physical, chemical and biological conditions (Beare *et al*., 1995, Ramette *et al*., 2007). Geochemical cycling rates might be dependent on the metabolite distances between different bacteria and Achaea in soils, yet, little is known about the spatial distribution of bacteria in soils and the drivers of the spatial heterogeneity. Several studies have focused on soil and sediment bacteria in order to understand spatial distribution of bacteria at different spatial scale level from the micro (Grundmann *et al*., 2004, Vos *et al*., 2013, Remenant *et al*., 2009, Córdova-Kreylos *et al*., 2006, Piza *et al*., 2005) to the macro scale (Fierer *et al*., 2006, Griffiths *et al*., 2009). Part of the difficulty is the physical-chemical description of the different samples at the micro-scale. The use of sediments to study spatial scale influences on microbial community distribution simplifies the micro-heterogeneity of physico-chemical parameters due to sediment pore water which increases local diffusion compared to unsaturated soils (Urban *et al*., 1997). Some studies have evaluated the bacterial composition and spatial distribution in sediment of brackish lakes (Thureborn *et al.*, 2013, Zaitseva *et al.*, 2014, Webster *et al.*, 2015, Pramanik et al., 2015). One marine sediment study has identified large scale shift in communities structure in water along a physico-chemical gradient, that of salinity, in the Baltic Sea (Herlemann *et al.*, 2011). While salinity clearly has an influence on microbial

community composition, several other physical-chemical characteristics also probably play important roles. Some of these characteristics could be driven by global land use practices. We hypothesized here that the spatial distribution of the microbial community could be modeled in part by sample characteristics and by global land use as determined using satellite data and geographic information systems (GIS). As lake sediments have less heterogeneity of abiotic conditions as compared to river sediments (Brönmark and Hansson, 2005) and are the result of the indigenous production and the transport of terrestrial material (organic or inorganic compound) from the drainage area (Beaulne *et al*., 2012, Teisserenc *et al.*, 2010), our sub-hypothesis is that the physico-chemical parameter having a spatial gradient at the scale of the lake (salinity in this case) will drive bacterial composition and diversity. The combination of marine and freshwater, gives Chilika Lake a wide range of habitat to sustain high diversity of microorganisms. We will try to identify marine and riverine influences using an approach that combines metagenomic (16S gene pyrotag of the V4-V6 region), GIS (Geographic Information Systems) and multivariate analysis.

Material and Methods

Study area

Chilika is the second largest brackish lagoon in the world (after Maracaibo Lake) and the largest in Asia situated on the east coast of India (between 19º28' and 19º54' North latitude and 85º05' and 85º 38' East longitude). Since 1981 the site is a designated first Indian wetland of international importance (Ramsar Site). The size of the lagoon fluctuates significantly during a year between 906 km² (summer) to 1165 km² (monsoon) Hydrological pattern of Chilika are impacted by three subsystems, namely 1) Mahanadi distributaries, 2) Streams of the western catchment which bring

in freshwater flows to the lake and 3) The Bay of Bengal which contributes highly saline sea-water. The lagoon is a collection of very shallow marine, brackish and freshwater ecosystems. The whole area is categorically organized into four sectors. The northern sector which receives direct discharge of freshwater from Mahanadi Delta, the central sector intermixing zone which is brackish, the southern sector which is observed to have higher salinity levels as compared to central sector and the outer channel which exchanges water between lagoon and the sea. Owing to a high salinity gradient the lake hosts a wide range of biodiversity in terms of macro organisms: 314 species of fish, 224 species of water birds and 729 angiosperms. That high biodiversity sustain an important population of fishermen estimated at about 140000 individuals (Kumar and Pattanaik, 2012).

After a decrease of the fishing productivity in the 1990 decade, a monitoring of the lake (water quality) has been done since 2000 and major management (reopening of the lake mouth at the outer channel) has been realized. The data produced by monitoring by Chilika development authority (CDA), have not only shown geographical gradient of physico chemical parameters, especially salinity, but also temporal changes between monsoon and dry season. They also identified 4 different hydrological zones. It i's a system to study the link between environmental parameters and composition structure of the bacterial communities inhabiting the sediment. A recent study (Delmont *et al.,* 2011) has shown, comparing the functional metagenome signature of different environment type, that Chilika Lake sediment have a different functional fingerprint than other environment, as different that soil, marine or human environment (Figure 1 –PCA Delmont).

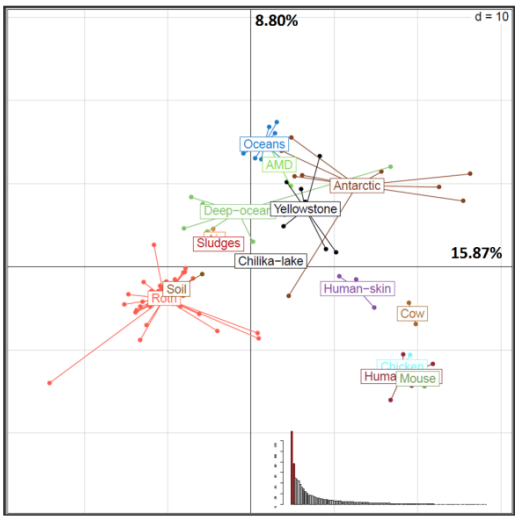


Figure 1: Functional fingerprints of diverse environment type

Physico chemical analysis

The monitoring of the sediment and water was done on 24 stations covering the whole lake (Figure 2 – monitoring station) in summer (March to June, 2011), rainy (July to October, 2011) and winter season (November, 2011 to February, 2012). Sediment and water samples were collected from stations spanning all four ecological sectors; northern, central, southern, and outer channel (Figure 2). Bottom surface sediment samples were collected through a Van Veen type of grab sampler (KC Denmark,) from each station during summer, rainy and winter seasons. Both sediment and water (water column) samples were immediately transferred to sterile plastic bottles of 500 ml and transported to the laboratory on ice and stored at 4^0C. Sediment samples were partially dried to improve lysis efficiencyand homogenized manually with the help of a sterile mortar and pestle inside a laminar flow hood. For each station pH, conductivity (mS/cm), turbidity (NTU), salinity (parts per thousand; ppt) and temperature of air and water (^{o}C) were measured onboard immediately with the help of a water quality probe analyzer (TOA DKK 24, Japan), depth and transparency (cm) of the water column were measured with the help of a measuring

tape and Secchi disk, respectively. Biological oxygen demand (ppm) and dissolved oxygen (ppm) were measured with an automated portable probe analyzer (VSI 07, VSI Electronics, Mohali, India).Concentration of nitrite, nitrate, ammonia, phosphate, silicate and total iron were measured on the collected water and total organic carbon on sediment sample (APHA, 1998). In order to see seasonal effect, we sampled each station 3 times, during summer, winter and rainy seasons.

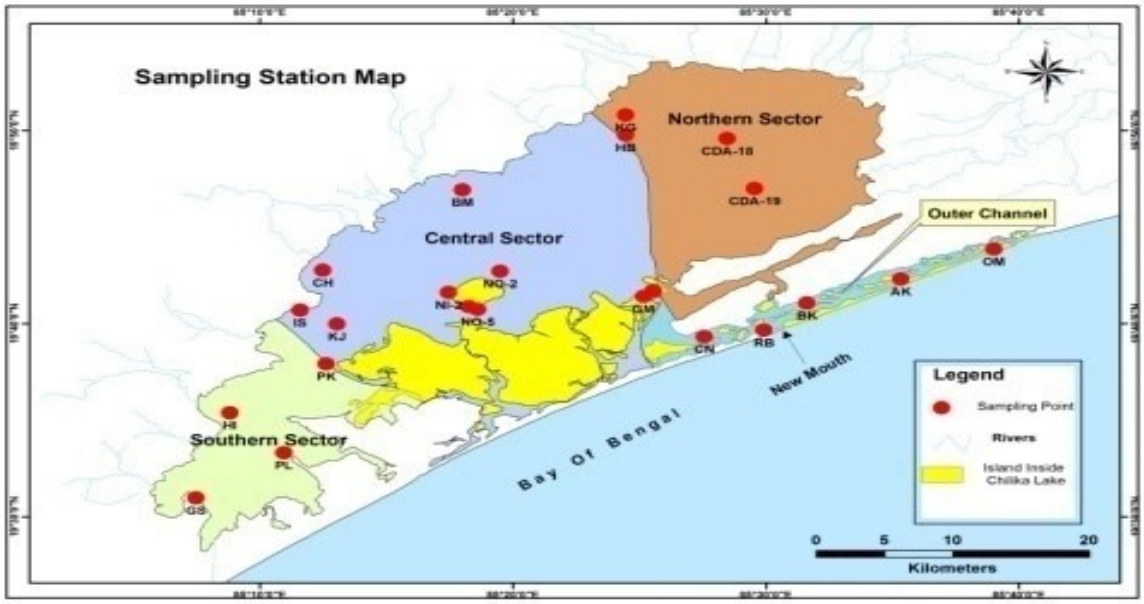


Figure 2: Grid map of Chilika Lake showing the position of 24 sampling stations spanning four different sectors, northern, central, southern, and outer channel.

GIS analysis

The proximal watershed limits were calculated using Digital Elevation Model (DEM) from SRTM images (NASA), and have been analyzed using GRASS (Geographical Resources Analysis Support System) an open-source GIS software (Beaulne 2008). Hydrological modeling was also performed in order to create stream, slope and

drainage layers (GRASS). To correspond to the scale level of the sediment and water sampling, we merged the sub-watershed in 4 large watersheds.

Land cover classes were assigned using four matricial Landsat TM 7 satellite images (july 2003) with a 27m resolution. Land cover classifications were then determined with the unsupervised maximum likelihood analysis method with ArcGIS software. We calculated the percentage of forested area, cultivated land, urban area, bare soils and wetland for each of the created watershed (Figure 3). We also hand draw areas of aquaculture (mainly shrimps and fish landing centers), from satellite images (CNES 2012 distribution Astrium services spot image**)**, known to be important in Chilika Lake and problematic in terms of water quality in other similar environment (Pushparajan and Soundarapandian 2010).

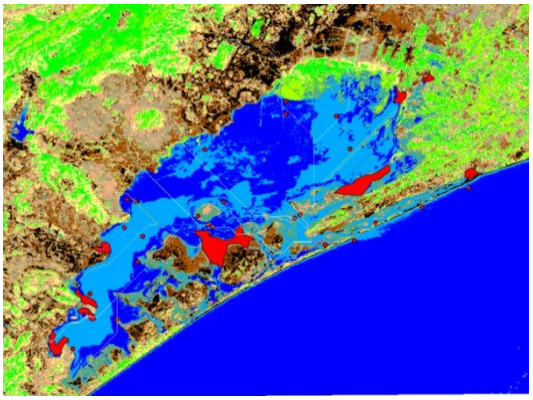


Figure 3: Interpolation of satellite images. Hand draw polygones of shrimp farming zone and fish landing area (red areas)

Interpolations with GRASS software were performed for the physico-chemical parameters measured for each sampled point and each season in order to find geographical and temporal gradient of distribution.

DNA analysis

Total DNA were extracted from 0.5g of sediment sample in duplicate using FastDNA® *Spin* Kit for soil DNA extraction (MP Biomedicals, Santa Ana, CA, USA) as per the manufacturer's protocol. DNA was extracted in duplicates for each sampling station and pooled to one tube per individual station and season. Quality and quantity of metagenomic DNA were verified by 1.2% of agarose gel (0.5XTAE) and Nanodrop (Epoch, BioTek Instruments, USA) analysis (A_{260}/A_{280}). DNA from a total of 72 samples from the 24 station (3 seasons) from 2011 and 2012 and 12 more samples after the "Phailin", a very severe cyclonicstorm (VSCS) made a landfall near 19.26^0 N / 84.82^0 E (near Gopalpur, Odisha) at 1600 UTC on 12th October, 2013 were extracted for bacterial diversity and community composition.

Total DNA extracted from the sediments was amplified for 16S rRNA genes by PCR. The hypervariable (V4-V6) regions of the 16S rRNA genes was amplified from the using bacterial primer pair 515F and 1061R (16S-0515F 5'-TGYCAGCMGCCGCGGTA-3' 16S-1061R 5'-TCACGRCACGAGCTGACG-3' ~560bp V4-V6 region) (Ref). PCR was carried out with the modified primers containing an adapter and a barcode sequence (Schloss *et al*., 2009). Paired end 454 pyrosequencing was performed on the GS-FLX 454 Titanium platform.

Analysis of the community composition

The reads were analyzed through the Quantitative Insights Into Microbial Ecology (QIIME) pipeline (http://qiime.org) and taxonomic identity of each read was assigned using the Bayesian rRNA classifier at 80% confidence threshold and QIIME (Wang *et al*., 2007; Cole *et al*., 2009, Caporaso *et al*., 2010). Sequences which could not be

classified to at least a kingdom level were excluded from subsequent analysis. For phylotyping of 16S rRNA data, sequences were aligned using the QIIME pyrosequencing pipeline. Operational taxonomic units (OTUs) were determined at phylum (80%), class (90%), order (92%), genus (95%), and species (97%) level.

Statistical and multivariate analysis

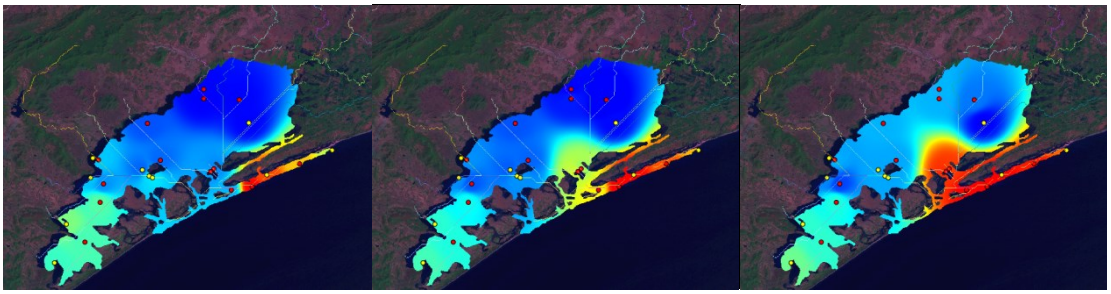
Statistical analyses were done using the JMP 11 software. In order to find link between physico-chemical parameters and relatives abundance of OTU's, a large set of linear relation were produced at different taxonomic level. To find similarities of community composition between samples, CANOCO 4.5 were used to draw Principal Component Analysis (PCA). Canonical Correspondence Analysis (CCA) was also performed to find, in the environmental dataset, which variables were driving the separation of the samples in term of community composition.

Results

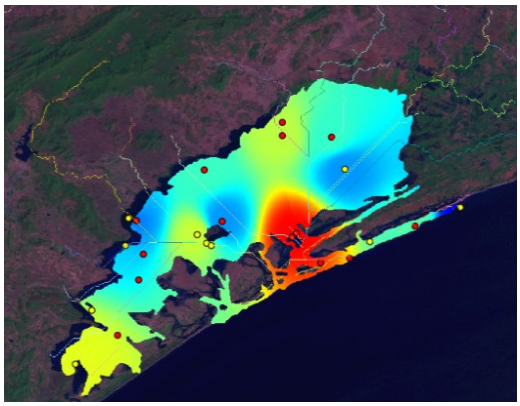
Physico-chemical distribution in the lake and spatial distribution of bacteria

In order to find correlation with physico-chemical parameter and the bacterial spatial distribution, we first looked at the spatial distribution of the physico-chemical parameter. The distribution of pH was between 6.9 and 9.5. We found no spatial patterns in the distribution. Only one sampled point (Kalupadaghat, KG) had a pH under 7.8 and only during the rainy season (8.6 during summer for the same point),

probably due to the increase of fresh water (lower pH) coming from the watershed during the monsoon. pH has been described to have strong effect on regulation of life processes and nutrient availability in aquatic systems. In the case of Chilika Lake, or at the scale of our study, pH did not appear to be a strong driver of the spatial distribution of bacteria and diversity. Turbidity was in a range between 1,1(NTU) to 171, but again non linear geographical gradient (patchy distribution), mainly depending on the seasonal effect. Highest turbidity was found in almost all post Phaillin samples where cyclonic storm causes mixing of bottom sediment layers. Salinity and conductivity were, non-surprisingly, correlated together. We focus on the salinity as it known to be a strong regulator for distribution and diversity of life. The range of salinity measured was between 0.1 and 32.3 ppt (g/kg). We observed a strong temporal variation and also a spatial gradient in the distribution. We created interpolated maps of the geographical variability of the salinity measured for each station (Figure 4 a b c). Nitrite, nitrate, ammonia, phosphate, silicate, total iron and total organic carbon were highly variable between sample site but also in the same sample site after the seasonal resampling, suggesting that there spatial heterogeneity occur at smaller scale. Moosoon sample had higher concentration of nutrient due to higher flux coming from watershed. It was impossible to link them with the variability of the community structure. The concentration of Nitrite (0.04 to 8.6 µmol/l), Nitrate (0.52 to 25.67 µmol/l) and Ammonia (7.31 to 206.41µmol/l) were low and had very high variability, improving the hypothesis of measurable effect on the micro-organisms distribution only at the smaller scale, sampling nest in that study was too large to identified trend.



a) Rainy Salinity b) Winter Salinity c) Summer Salinity



d) Interpolation of bacterial diversity (Shannon index)

Figure 4: Interpolation of the salinity measured in Rainy (a), Winter (b) and Summer (c) season. Interpolation of the Shannon index(d). Colour gradient from red (highest concentration) to blue (lowest concentration)

GIS analysis

Interpolation of physico-chemical parameters were done for each season. Only salinity showed a spatial gradient in the distribution. Interpolation of the diversity (Shannon index at genus level) was also done using GRASS software. We observed a significant “hot spot” for bacterial diversity. Comparing the geographical zone of higher diversity with the three maps of the salinity we find a relation between the geographical areas where the salinity was highly variable during the year (seasonal effect, Figures 4a,b,c) and the geographical area where we measured the highest diversity (Figure 4d). The 4 sampled point, Godhimukh (GM), Magarmukh (MM),

Chakanasi (CN) and Rambhartia (RB), where salinity always stay high during three season had much lower Shannon index, suggesting that it is not an effect of the salinity but an effect of the continuous mixing of water from different watershed and changing environment from low to high salinity.

Statistical analysis

We performed multivariate analysis (PCA, CCA) with CANOCO v4.5, to find the main driver of the variability in our samples. We observed a separation of the samples following principally the variation of the salinity and conductivity (Figure 5). The spatial effect can also be responsible for the differences in community structure as all the samples with the higher salinity are located in the same geographical zone, the outer channel. A PCA was also produced to see the seasonal effect on the bacterial community structure. Considering all the samples, we were not able to see any evidence of major shift in the community structure related with the seasonal effect (Figure 6).

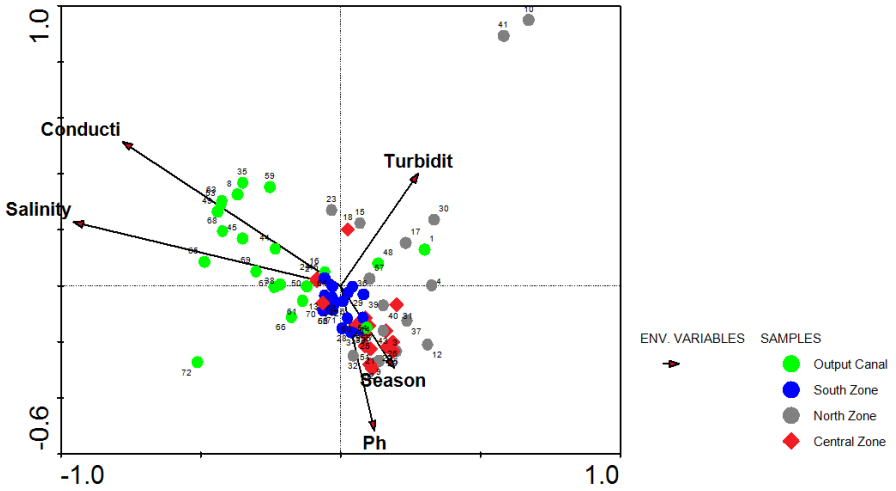


Figure 5: PCA at the genus level (72 sample), environmental drivers and zonal effect.

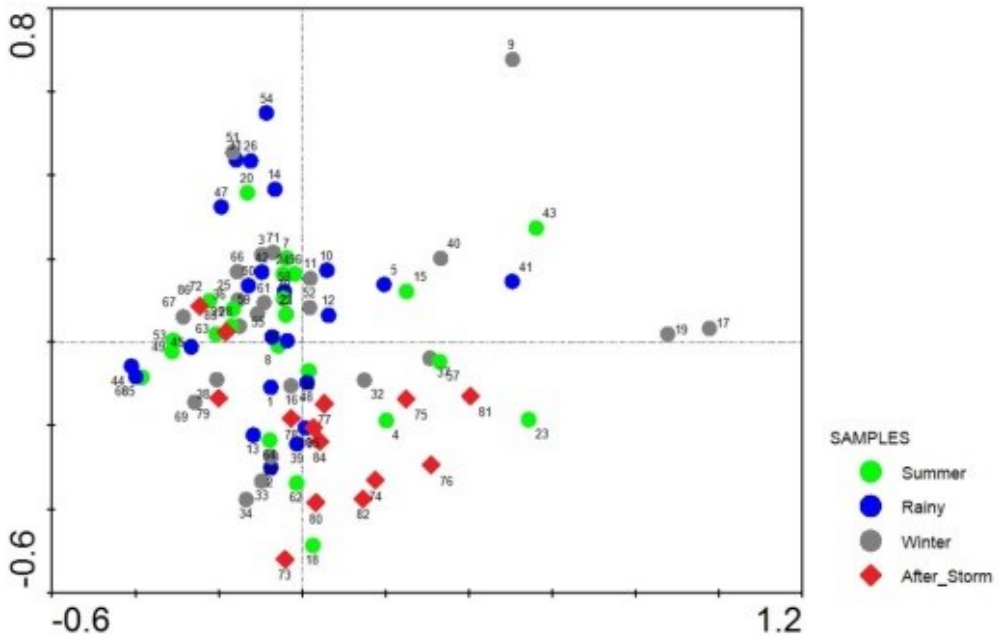


Figure 6: PCA on the genus level (72 samples) and seasonal effect. No significant shift in the community structure

In order to observe significant deviations in the bacterial community structure, twelve samples were collected just after a major storm in October, 2013. We plotted post-storm (30 October-25 November 2013) and pre-storm samples (March 2011-February 2012) on a PCA to visualize if there are differences in the community structure of those two groups. The dark dot (green before cyclone and red after cyclone) are the one with high salinity and the light (green before cyclone and red after cyclone), the one located in low salinity zone. The samples with high salinity before (dark green) and after the storm (dark red) shown a similar community structure (Figure 7). The low salinity samples taken after the storm (light red) have different community structure than the one sampled before the storm (light green). We compared the same twelve sample points after and before the storm (Figure 8), we can see a little shift in the community structure probably due to the mixing of the sediment (sediment sampled were at a depth between 1 m and 3.2 m) and a major

influx of matters from the watershed. The sample points located near by the ocean were not significantly different after the storm, probably used to be more impacted by the effect of tide.

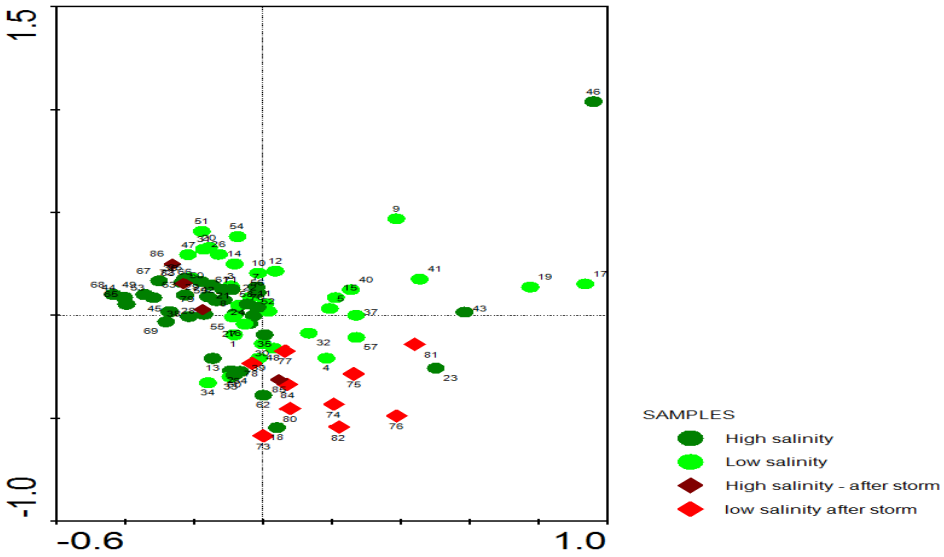


Figure 7:PCA at the genus level (72 sample), high and low salinity vs before and after the storm.

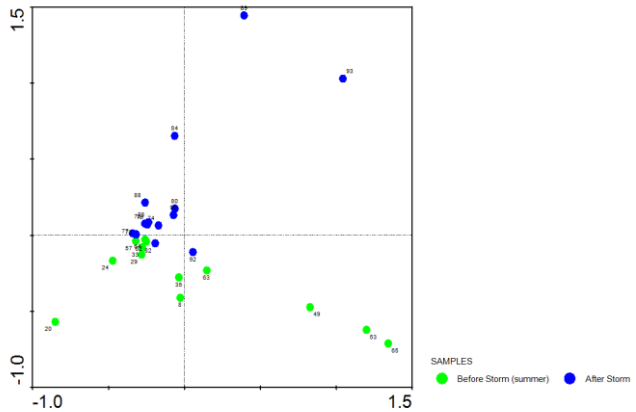


Figure 8: PCA at the genus level (24 samples), before and after the storm.

Analysis of the communities structure

The number of reads per samples were between 5 387 to 38 265. Taxonomy was done using QIIME pipeline. The most dominant phyla in our sample were Proteobacteria with a relative abundance up to 0.79 and never under 0.37 (average 0.573). The second more abundant phyla were the Chloroflexi (0.01 to 0.24, average 0.086). In order, the next phyla were Bacteriodetes (0.006 to 0.17, average 0.047), Planctomycetes (0.001 to 0.08, average 0.038), Acidobacteria (0.002 to 0.07, average 0.031), Firmicutes (0.0004 to 0.15, average 0.028), Gemmatimonadetes (0.002 to 0.08, average 0.027), Actinobacteria (0.004 to 0.08, average 0.023), Verrucomicrobia (0.0004 to 0.014, average 0.021) and Nitrospirae (0.0005 to 0.06, average, 0.017) (Figure 9). Looking more specifically at Proteobacteria (largely dominant phyla in all the samples), we have a specific distribution at the class level. Interestingly the most dominant class of Proteobacteria was Gammaproteobacteria with an average of 0.252 of relative abundance, followed by Deltaproteobacteria (0.178) and with a much lower abundance, Betaproteobacteria (0.066), Alphaproteobacteria (0.045) and Epsilonproteobacteria (0.026). The really high abundance of Gammaproteobacteria was mainly due to 4 order, Chromatiales (up to 0.19), Xanthomonadales (up to 0.12), Pseudomonadales (up to 0.21) and Thiotrichales (up to 0.21). Some family and genus have also been measured to be in high concentration in some of our samples like, Piscirickettsiaceae (up to 0.21), Pseudomonas (up to 0.09). Surprisingly, the relative abundance of Acidobacteria was observed, considering the high pH in almost every part of the lake. In contrary with the relative abundance of Acidobacteria that have been described to be correlated with soils pH in studies (Griffith 2011, Fierer 2006), we have not observed

any correlation with the pH and the relative abundance of the phyla Acidobacteria (Figure 10) and also with the different genera composing the phyla.

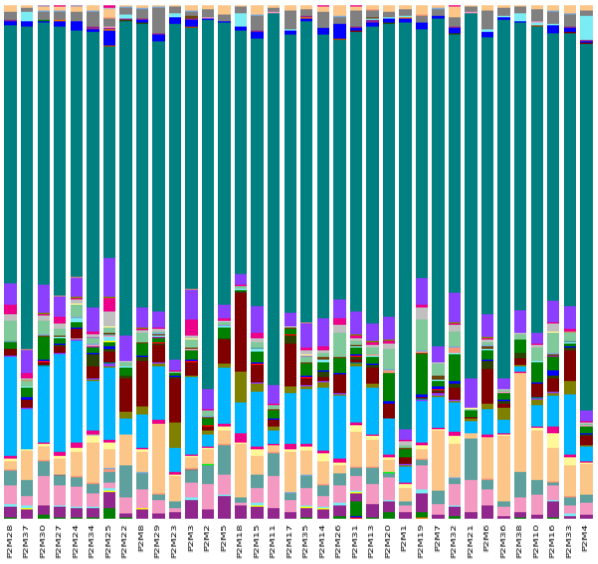


Figure 9: Distribution of Phyla (QIIME)

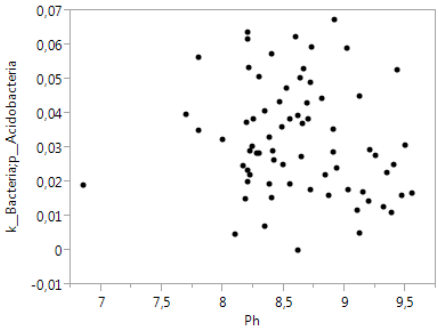


Figure 10: Relation between Acidobacteria abundance and pH

We also saw a negative effect of the salinity on the diversity (exception of the point located in the mixing zone) (Figure 11a). We have a positive relation between the salinity and the relative abundance of Proteobacteria (Figure 11b). As the relative

abundance of Proteobacteria is really high we also saw a strong negative relation with the abundance of Proteobacteria and the diversity (Figure 11c)

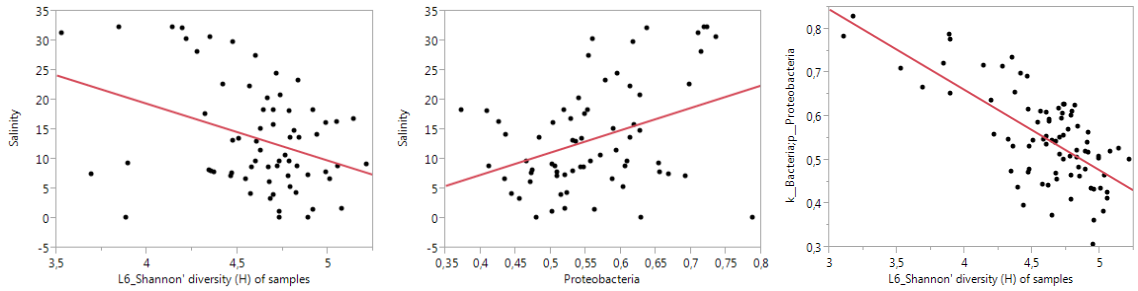


Figure 11:a) Effect of salinity on the diversity (Shannon Index) b) Effect of salinity on Proteobacteria relative abundance (R^2 0.1896 , $p<0.0002$) c) Effect of the relative abundance of Proteobacteria and diversity (Shannon index)

Link between physico-chemical variable and bacterial community structure have been shown with the distribution of the most relatively abundant Phylum and Class of bacteria found in the Chilika Lake sediment, Proteobacteria and Gammaproteobacteria, and the distribution of salinity. With a R^2 of 0.2829 ($p<0.0001$) (Figure 12), salinity is a strong predictor of the distribution of the Gammaproteobacteria in the sediment of the Chilika Lake. At the Phylum level less stronger relation (R^2 0.1896 , $p<0.0002$) (Figure 11b) was observed with salinity, mainly due to the dominant class of the Phyla, Gamaproteobacteria (Figure 12).

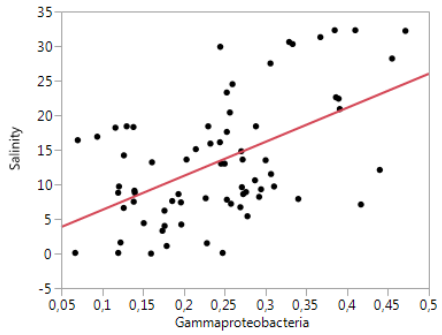


Figure 12: Effect of salinity on Gammaproteobactria relative abundance, R^2 of 0,2829 $p<0,0001$

Discussion

Drivers of the community structure

Research has shown that ecological changes in the Chilika lake system and its fisheries for several years were the effect of water exchange between the lake and the sea that is responsible for control of the salinity, siltation, macrophyte infestation and increase of marine forms (Ghosh *et al*., 2006; Satyanarayana *et al*., 1999; Sahu *et al*., 2014). The community structure and distribution in Chlika Lake observed in this study is similar to what have been observed in 16S gene study on brackish water, Baltic sea (Herlemann, 2011), in sediment of Sundarbans mangrove wetland (Basak *et al*., 2014), in Chilika Lake (Pramanik *et al*., 2015) and in Brazilian mangrove (Andreote *et al*., 2012). Salinity was the only variable that we identified a spatial gradient of distribution, at the scale of our sampling, in the lagoon. The distributions of other variables measured in the study were having a much higher or lower spatial scale of variation. The higher scale of variation will affect all the samples, making it impossible to measure their effect on the spatial distribution of the bacteria in the sediment (climatic conditions, watershed land cover). On the other hand the distribution and variability of nitrite, nitrate, ammonia, phosphate, silicate, total iron and total organic carbon occurred at a much smaller scale than the scale of our sample. Hot spot for those variables are visible at the micro scale (Parkin 1987, Schramm 1999). The size of samples (0.5g), the nesting of the sampling effort and the fact that we are not using the exact same sample to do chemical analysis and 16S analysis make it almost impossible to find correlation with those micro scale variables and to identify drivers of the community structure at the scale of our study (Figure 13).

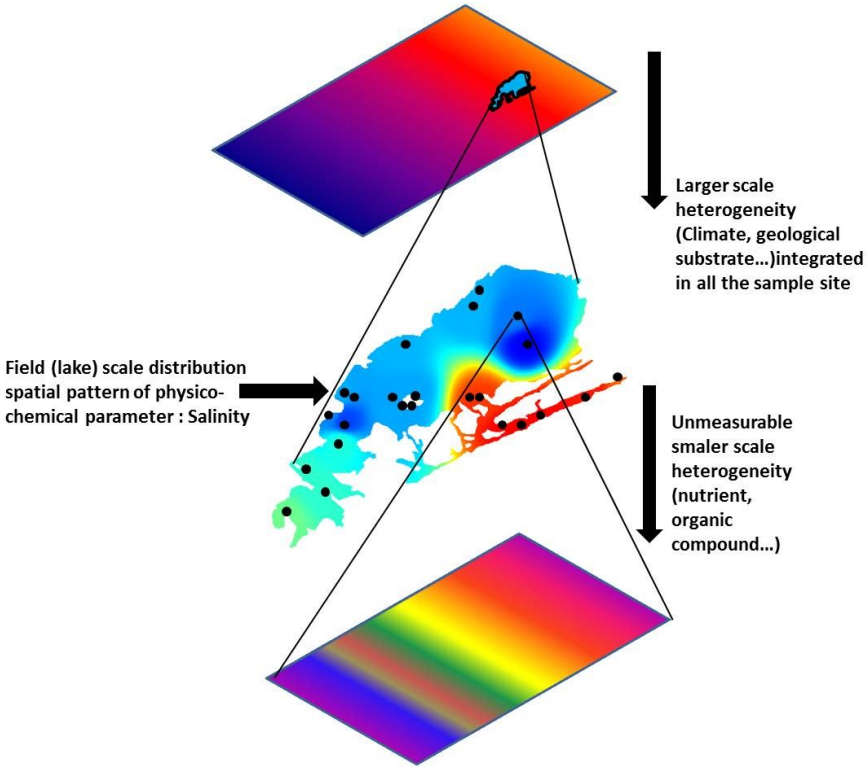


Figure 13: Multiple scale heterogeneity of the spatial distribution of variables driving the composition of bacterial communities

Another limitation in the analysis of the spatial distribution of bacteria and the identification of the physico-chemical drivers is the importance of the transport of material in the lake. We identified all the samples, where we had really high concentration (more than 0.08 relative abundance) of a family (Chromatiales, Xanthomonadales, Pseudomonadales and Thiotrichales) or a genus (Piscirickettsiaceae and Pseudomonas, more than 0.05 relative abundance). Coupling a map of the areas of aqua farming and a stream map (produced with Hydrological Modeling, GRASS) we observed that almost all the point with high concentration of Gammaproteobacteria single family or genus were located on a stream coming from aqua farming areas (Figure 14). Those four families and two

genera are known to contain pathogen species; some of them have been associated with fish pathogen in salmon aqua farming (Austin & Austin 2007, Birkbeck 2011). The spatial analysis also highlighted hot spot of bacterial diversity where we have high seasonal variation of the salinity. The change in salinity is mainly due to the huge amount of freshwater coming in the lagoon during the monsoon. On a daily basis some changes of salinity can be due to the tidal effect.

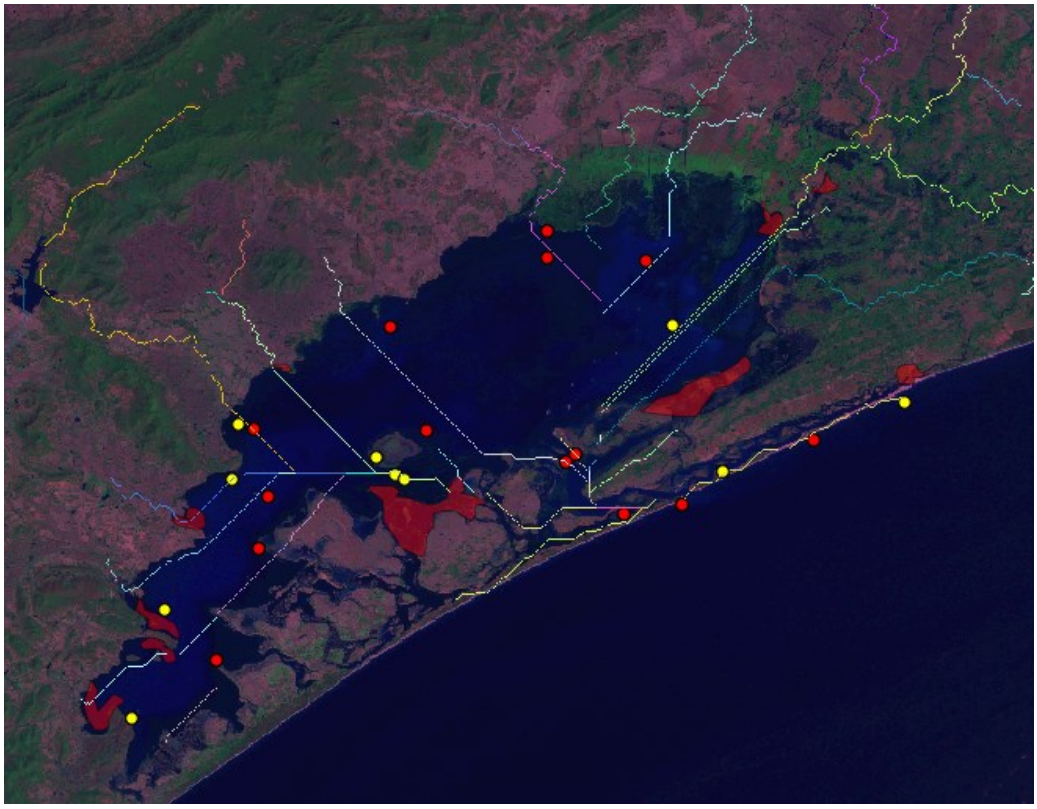


Figure 14: Stream analysis, high abundance of single genus or family (yellow dot) and fish farms (red polygons).

Conclusion

The first description of the spatial distribution of bacteria in the Chilika Lake sediment have highlighted the predominance of Gammaproteobacteria. We were able to identify trends in the spatial and temporal distribution of that class in that particular environment. The identification of hot spot of some family or genus of that class link with transport of material in the lake (through stream) from fish farming areas, demonstrate the importance of spatial analysis. The only variable, salinity, that had a spatial pattern of distribution at the same scale level as the sampling effort done for this study, was the only one that we have been able to correlate with the community structure of our sample. That strong relation between salinity and relative abundance of Gamaproteobacteria, comfort our hypothesis that the spatial distribution of bacteria in sediment is driven by the physico-chemical parameter and occur at the same spatial scale level. The importance of the temporal dimension has been demonstrated with the identification of hot spot of diversity, where we have temporal change of salinity. Other questions need to be answered regarding the spatial distribution of bacteria in Chilika Lake, how does the presence of Phragmites in some area affects the bacterial population? How does the composition is influenced in Nalaban island where lot of nitrogen and phosphorus inputs are added by the birds.

Using only statistical approach to identify physicochemical drivers of the community structure can be meaningless, if we are not considering the spatial and the temporal scale. The variables having a really low concentration can have hot spot and create specific niches for bacteria at the micro scale. But, comparing samples of 0.5g will not allow us to measure the importance of that variability in the whole community of bacteria present in that size of sample. On another, the variability occurring at a higher spatial scale level will be integrated in all the samples.

Acknowledgement

We are grateful to the Chilika Development Authority (CDA), Bhubaneswar and World Bank-supported Integrated Coastal Zone Management Project for providing financial assistance. Staff of the CDA are acknowledged for their help and providing necessary facilities in collection of sediment samples.

Bibliography

1. Andreote, F. D., Jiménez, D. J., Chaves, D., Dias, A. C. F., Luvizotto, D. M., Dini-Andreote, F., & de Melo, I. S. (2012). The microbiome of Brazilian mangrove sediments as revealed by metagenomics. *PLoS One*, *7*(6), e38600.

2. Austin, B., & Austin, D. A. (2007). *Bacterial fish pathogens: disease of farmed and wild fish*. Springer Science & Business Media.

3. Basak, P., Majumder, N. S., Nag, S., Bhattacharyya, A., Roy, D., Chakraborty, A., ... & Bhattacharyya, M. (2015). Spatiotemporal analysis of bacterial diversity in sediments of Sundarbans using parallel 16S rRNA gene tag sequencing. *Microbial ecology*, *69*(3), 500-511.

4. Beaulne, J. S. (2008). Modélisation de la présence de mercure dans la chair des brochets et des dorés des lacs les plus pêchés de l'Abitibi-Témiscamingue: une approche par les systèmes d'information géographique. Master Thesis Report, UQAM

5. Beaulne, J. S., Lucotte, M., Paquet, S., & Canuel, R. (2012). Modeling mercury concentrations in northern pikes and walleyes from frequently fished lakes of Abitibi-Témiscamingue (Quebec, Canada): a GIS approach. *Boreal environment research*, *17*(3/4), 277-290.

6. Beare MH, Coleman DC, Crossley Jr DA, Hendrix PF, Odum EP (1995). A hierarchical approach to evaluating the significance of soil biodiversity to biogeochemical cycling. *Plant and Soil*, 170: 5-22.

7. Birkbeck, T. H., Feist, S. W., & Verner–Jeffreys, D. W. (2011). Francisella infections in fish and shellfish. *Journal of fish diseases*, *34*(3), 173-187.

8. Brönmark, C., & Hansson, L. A. (2005). *The biology of lakes and ponds*. Oxford University Press.

9. Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, *7*(5), 335-336.

10. Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., ... & Tiedje, J. M. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic acids research*, *37*(suppl 1), D141-D145.

11. Córdova-Kreylos, A. L., Cao, Y., Green, P. G., Hwang, H. M., Kuivila, K. M., LaMontagne, M. G., & Scow, K. M. (2006). Diversity, composition, and geographical distribution of microbial communities in California salt marsh sediments. *Applied and environmental microbiology*, *72*(5), 3357-3366.

12. Delmont, T. O., Malandain, C., Prestat, E., Larose, C., Monier, J. M., Simonet, P., & Vogel, T. M. (2011). Metagenomic mining for microbiologists. *The ISME journal*, *5*(12), 1837-1843.

13. Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 626-631.

14. Ghosh, A. K., Pattnaik, A. K., & Ballatore, T. J. (2006). Chilika Lagoon: Restoring ecological balance and livelihoods through re-salinization. *Lakes & Reservoirs: Research & Management*, *11*(4), 239-255.

15. Griffiths G., Thomson R.I., James B.C., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of British soils. *Environmental microbiology*, *13*(6), 1642-1654.

16. Grundmann, G. L. (2004). Spatial scales of soil bacterial diversity–the size of a clone. *FEMS Microbiology Ecology*, *48*(2), 119-127.

17. Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., & Andersson, A. F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME journal*, *5*(10), 1571-1579.

18. Kumar, R., & Pattnaik, A. K. (2010). Integrated planning management for Chilika Lake. *Chilika Newslett*, *5*, 1-12.

19. Parkin, T. B. (1987). Soil microsites as a source of denitrification variability. *Soil Science Society of America Journal*, *51*(5), 1194-1199.

20. Piza, F. F., Prado, P. I., & Manfio, G. P. (2005). Investigation of bacterial diversity in Brazilian tropical estuarine sediments reveals high actinobacterial diversity. *Antonie Van Leeuwenhoek*, *86*(4), 317-328.

21. Pramanik, A., Basak, P., Banerjee, S., Sengupta, S., Chattopadhyay, D., & Bhattacharyya, M. (2015). Pyrosequencing based profiling of the bacterial community in the Chilika lake, the largest lagoon of India. *Genomics Data*, *4*, 112-114.

22. Pushparajan, N., & Soundarapandian, P. (2010). Recent farming of marine black tiger shrimp, Penaeus Monodon (Fabricius) in South India. *African Journal of Basic & Applied Sciences*, *2*(1-2), 33-36.

23. Remenant, B., Grundmann, G. L., & Jocteur-Monrozier, L. (2009). From the micro-scale to the habitat: assessment of soil bacterial community structure as shown by soil structure directed sampling. *Soil Biology and Biochemistry*,*41*(1), 29-36.

24. Schramm, A., de Beer, D., van den Heuvel, J. C., Ottengraf, S., & Amann, R. (1999). Microscale Distribution of Populations and Activities ofNitrosospira and Nitrospira spp. along a Macroscale Gradient in a Nitrifying Bioreactor: Quantification by In Situ Hybridization and the Use of Microsensors. *Applied and environmental microbiology*, *65*(8), 3690-3696.

25. Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., ... & Weber, C. F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*,*75*(23), 7537-7541.

26. Teisserenc, R., Lucotte, M., Houel, S., & Carreau, J. (2010). Integrated transfers of terrigenous organic matter to lakes at their watershed level: A combined biomarker and GIS analysis. *Geochimica et Cosmochimica Acta*,*74*(22), 6375-6386.

27. Thureborn, P., Lundin, D., Plathan, J., Poole, A. M., Sjöberg, B. M., & Sjöling, S. (2013). a metagenomics transect into the deepest point of the Baltic Sea reveals clear stratification of microbial functional capacities. *PloS one*, *8*(9).

28. Urban, N. R., Dinkel, C., & Wehrli, B. (1997). Solute transfer across the sediment surface of a eutrophic lake: I. Porewater profiles from dialysis samplers. *Aquatic Sciences*, *59*(1), 1-25.

29. Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of bacterial diversity in soil. *FEMS microbiology reviews*, *37*(6), 936-954.

30. Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiology Ecology*, *62*(2), 142-160.

31. Sahu, B. K., Pati, P., & Panigrahy, R. C. (2014). Environmental conditions of Chilika Lake during pre and post hydrological intervention: an overview. *Journal of Coastal Conservation*, *18*(3), 285-297.

32. Satyanarayana, C. H. (1999). *Hydrographic and phytoplankton characteristics of Chilka Lake, a brackishwater lagoon, on the east coast of India* (Doctoral dissertation, Ph. D. thesis, Andhra University, Visakhapatnam, India).

33. Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*, *73*(16), 5261-5267.

34. Webster, G., O'Sullivan, L. A., Meng, Y., Williams, A. S., Sass, A. M., Watkins, A. J. & Weightman, A. J. (2015). Archaeal community diversity and abundance changes along a natural salinity gradient in estuarine sediments. *FEMS microbiology ecology*, *91*(2), 1-18.

35. Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences*, *95*(12), 6578-6583.

36. Zaitseva, S. V., Abidueva, E. Y., Namsaraev, B. B., Wang, L., & Wu, L. (2014). Microbial community of the bottom sediments of the brackish Lake Beloe (Transbaikal region). *Microbiology*, *83*(6), 861-868.

List of captions

Conclusion

The importance of understanding the spatial distribution of bacteria in space and the parameters driving that spatial distribution is critical. Microorganisms were the first form of life on earth and everything that came after were due to their activity and implication in global cycle. In a changing world, the impact of human activities on life has been well studied but little is known about microorganisms. In this study, we identified some trends at different spatial scales that tend to confirm our general hypothesis “Spatial distribution of bacteria in soils and sediments are driven by physico-chemical parameters, land cover and land use, and that spatial variability occur at the same spatial scales”.

As presented in chapter one, the size of soil metagenome depends on what we consider as a soil unit, the size of the sample and the deepness of the sequencing. Although soil metagenomes are most similar to other soil metagenomes among the different possible ecosystems, significant differences in term of community structure can be observed at multiple scales. At the micro-scale, the size of bacterial environment, few parameters have been identified to understand how bacteria are organized in space. Some studies have identified micro-niches, but not necessarily the structure of communities associated with those niches. We know that bacteria favor pores filled with air or water and different types of communities inhabit those different niches. Sizes of aggregates have been studied in detail as we can collect multiple aggregates of the same size to sequence them. Correlations between spatial variability of physico-chemical parameter (mainly type of organic matter) and the size of aggregates seems to drive the community structure inhabiting these aggregates. At the field scale, the variables that drive the bacterial community structure are

mainly vegetation cover and nutriment distribution. Again, and specifically in agricultural fields, the only the variables having a spatial heterogeneity at the field-scale can be use to find correlations with the relative abundance of some taxa. At a larger scale (regional to global scale), few variables have been identified that drive the community structure. The only variables demonstrating a pattern of distribution (patch, increasing, decreasing) can be used to model the spatial distribution of bacteria.

The question “Can a variable with a spatial distribution at a larger scale (cm) overcome the micro-scale heterogeneity?” To answer that question, we induce and important chemical change on a core of soil by adding diesel on half of the core. After 14 days of incubation of the contaminated and uncontaminated part of the core, we were able to test our hypothesis “if the scale of the perturbation corresponds to the scale of the sampling, we should observe a shift in organisms that are adapted to the new condition and overcome a part of the smaller scale heterogeneity”. We observed a significative increase in taxa known to be present in hydrocarbon-contaminated soils. Furthermore, we need to consider soils as a 3-dimensional environment as we found higher vertical differences then horizontal differences in term of community composition, and the effect of contamination was attenuated in the deeper samples.

For environmental studies, we need to consider much larger areas to measure the impact of changes in the environment on global biogeochemical cycle, but since physico-chemical parameters are difficult to model at large scales, the question “Do some large scale bioindicator can integrate groups of variables to model the distribution of bacteria for an entire region” needs to be addressed. By coupling

phylogenetic microarrays, physical and chemical analysis, multivariate analysis tools and large-scale geographical information system (GIS) analysis, we compared the influence of physico-chemical parameters, alone or in group, with land cover analysis. The example of pH (single variable) versus forest (integrative variable) effects on relative abundances of Acidobacteria supported our hypothesis that variables that integrate holistic numbers of physico-chemical data (e.g., forests) can be better indicators of community structure than the physico-chemical data alone. This approach also demonstrated the feasibility of using GIS tools, and satellite and DEM images for large-scale spatial analyses of bacterial distribution by taking into account land cover and water/rain/river run-off.

Finally in chapter 4, we presented a field study on sediment of Chilika Lake (the first large scale and temporal description of the bacterial population in brackish lake sediment) in order to verified *in situ* the hypothesis “variables with geographical gradients at the scale of the sampling should be strong drivers of the community structure”. With a sampling strategy covering the gradient of salinity in the lake, we found a strong correlation between the salinity and the abundance of Proteobacteria and more specifically Gammaproteobacteria (the most abundant class). The use of spatial analysis tools also allow us to identified hot spots of diversity where we have sesonal changes of salinity. With simple statistical approaches, the variables, salinity and season, taken separately were not indicators of the microbial diversity. In addition, analyses of satellite images (fish farms) and hydrological modeling (hydraulic flows) identified hot spots of some members of Gammaproteobacteria known to be fish pathogens and associated with fish farming activities.

In this study, we highlighted the importance of considering multiple scales to understand the spatial distribution of bacteria in soils and sediments. By integrating 16S rRNA gene data and physico-chemical data in a GIS system, we were able to focus on variables having a pattern of spatial distribution compatible with the area of the samples we were analyzing. To go further, by superimposing multiple layers of different spatial scale drivers in a GIS system, we can move to a more global representation of the spatial distribution of bacteria and model the impact of changes on the environment.

Bibliography

Alvarez, V. M., Santos, S. C., Casella, R. C., Vital, R. L., Sebastin, G. V., & Seldin, L. (2008). Bioremediation potential of a tropical soil contaminated with a mixture of crude oil and production water. *Journal of microbiology and biotechnology*, *18*(12), 1966-1974.

Andreote, F. D., Jiménez, D. J., Chaves, D., Dias, A. C. F., Luvizotto, D. M., Dini-Andreote, F., & de Melo, I. S. (2012). The microbiome of Brazilian mangrove sediments as revealed by metagenomics. *PLoS One*, *7*(6), e38600.

Angel, R., Soares, M. I. M., Ungar, E. D., & Gillor, O. (2010). Biogeography of soil archaea and bacteria along a steep precipitation gradient. *The ISME journal*, *4*(4), 553-563.

Auffret, M. D., Yergeau, E., Labbé, D., Fayolle-Guichard, F., & Greer, C. W. (2015). Importance of Rhodococcus strains in a bacterial consortium degrading a mixture of hydrocarbons, gasoline, and diesel oil additives revealed by metatranscriptomic analysis. *Applied microbiology and biotechnology*, *99*(5), 2419-2430.

Austin, B., & Austin, D. A. (2007). *Bacterial fish pathogens: disease of farmed and wild fish*. Springer Science & Business Media.

Austrheim, G. (2002). Plant diversity patterns in semi-natural grasslands along an elevational gradient in southern Norway. *Plant Ecology*, *161*(2), 193-205.

Baas-Becking, Lourens G.M. (1934), *Geobiologie of inleiding tot de milieukunde*, The Hague, the Netherlands: W.P. Van Stockum & Zoon

Bach and Hofmockel, 2014, Soil aggregate isolation method affects measures of intra-aggregate extracellular enzyme activity. *Soil Biology and Biochemistry*, *69*, 54-62.

Bates, S. T., Berg-Lyons, D., Caporaso, J. G., Walters, W. A., Knight, R., & Fierer, N. (2011). Examining the global distribution of dominant archaeal populations in soil. *The ISME journal*, *5*(5), 908-917.

Basak, P., Majumder, N. S., Nag, S., Bhattacharyya, A., Roy, D., Chakraborty, A., ... & Bhattacharyya, M. (2015). Spatiotemporal analysis of bacterial diversity in sediments of Sundarbans using parallel 16S rRNA gene tag sequencing. *Microbial ecology*, *69*(3), 500-511.

Beaulne, J. S. (2008). Modélisation de la présence de mercure dans la chair des brochets et des dorés des lacs les plus pêchés de l'Abitibi-Témiscamingue: une approche par les systèmes d'information géographique. Master Thesis Report, UQAM

Beaulne, J. S., Lucotte, M., Paquet, S., & Canuel, R. (2012). Modeling mercury concentrations in northern pikes and walleyes from frequently fished lakes of Abitibi-Témiscamingue (Quebec, Canada): a GIS approach. *Boreal environment research*, *17*(3/4), 277-290.

Beare MH, Coleman DC, Crossley Jr DA, Hendrix PF, Odum EP (1995). A hierarchical approach to evaluating the significance of soil biodiversity to biogeochemical cycling. *Plant and Soil*, 170: 5-22.

Berg, G., & Smalla, K. (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS microbiology ecology*, *68*(1), 1-13.

Bik, H. M., Porazinska, D. L., Creer, S., Caporaso, J. G., Knight, R., & Thomas, W. K. (2012). Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in ecology & evolution*, *27*(4), 233-243.

Binkley, D., C.T. Driscoll, H.L. Allen, P. Shoeneberger, and D. McAvoy. (1989). Acidic Deposition and Forest Soils: Context and Case Studies in the Southwestern United State. Manual. Edwards Brothers Publication, Michigan.

Binkley, D., & Fisher, R. (2012). *Ecology and management of forest soils*. John Wiley & Sons.

Birkbeck, T. H., Feist, S. W., & Verner–Jeffreys, D. W. (2011). Francisella infections in fish and shellfish. *Journal of fish diseases*, *34*(3), 173-187.

Brönmark, C., & Hansson, L. A. (2005). *The biology of lakes and ponds*. Oxford University Press.

Bryant, J. A., Lamanna, C., Morlon, H., Kerkhoff, A. J., Enquist, B. J., & Green, J. L. (2008). Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences*, *105*(Supplement 1), 11505-11511.

Calderón, F. J., Reeves, J. B., Collins, H. P., & Paul, E. A. (2011). Chemical differences in soil organic matter fractions determined by diffuse-reflectance mid-infrared spectroscopy. *Soil Science Society of America Journal*, *75*(2), 568-579.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., ... & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, *7*(5), 335-336.

Castro, H. F., Classen, A. T., Austin, E. E., Norby, R. J., & Schadt, C. W. (2010). Soil microbial community responses to multiple experimental climate change drivers. *Applied and environmental microbiology*, *76*(4), 999-1007.

Chan, O. C., Casper, P., Sha, L. Q., Feng, Z. L., Fu, Y., Yang, X. D., ... & Zou, X. M. (2008). Vegetation cover of forest, shrub and pasture strongly influences soil

bacterial community structure as revealed by 16S rRNA gene T-RFLP analysis. *FEMS microbiology ecology*, *64*(3), 449-458.

Chen, C., Bai, Y. H., Liang, J. S., & Yuan, L. J. (2015). [Application of FISH-NanoSIMS technique in environmental microbial ecology study]. *Huan jing ke xue= Huanjing kexue/[bian ji, Zhongguo ke xue yuan huan jing ke xue wei yuan hui" Huan jing ke xue" bian ji wei yuan hui.]*, *36*(1), 244-251.

Chenu and Consantino, 2011, Microbial Regulation of Soil Structural Dynamic. *The Architecture and Biology of Soils*, pages 17-69

Chikere, C. B., Okpokwasili, G. C., & Chikere, B. O. (2009). Bacterial diversity in a tropical crude oil-polluted soil undergoing bioremediation. *African Journal of Biotechnology*, *8*(11).

Cho JC, Tiedje JM (2002). Quantitative detection of microbial genes by using DNA microarrays. *Appl Environ Microbiol.* 68: 1425-1430.

Chu, H., Fierer, N., Lauber, C. L., Caporaso, J. G., Knight, R., & Grogan, P. (2010). Soil bacterial diversity in the Arctic is not fundamentally different from that found in other biomes. *Environmental Microbiology*, *12*(11), 2998-3006.

Chu, H., Neufeld, J. D., Walker, V. K., & Grogan, P. (2011). The influence of vegetation type on the dominant soil bacteria, archaea, and fungi in a low Arctic tundra landscape. *Soil Science Society of America Journal*, *75*(5), 1756-1765.

Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., ... & Tiedje, J. M. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic acids research*, *37*(suppl 1), D141-D145.

Córdova-Kreylos, A. L., Cao, Y., Green, P. G., Hwang, H. M., Kuivila, K. M., LaMontagne, M. G., & Scow, K. M. (2006). Diversity, composition, and geographical distribution of microbial communities in California salt marsh sediments. *Applied and environmental microbiology*, *72*(5), 3357-3366.

Curtis, T. P., Sloan, W. T., & Scannell, J. W. (2002). Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences*,*99*(16), 10494-10499.

Davinic, M., Fultz, L. M., Acosta-Martinez, V., Calderón, F. J., Cox, S. B., Dowd, S. E., & Moore-Kucera, J. (2012). Pyrosequencing and mid-infrared spectroscopy reveal distinct aggregate stratification of soil bacterial communities and organic matter composition. *Soil Biology and Biochemistry*,*46*, 63-72.

Delmont, T. O., Malandain, C., Prestat, E., Larose, C., Monier, J. M., Simonet, P., & Vogel, T. M. (2011). Metagenomic mining for microbiologists. *The ISME journal*, *5*(12), 1837-1843.

Delmont, T. O., Robe, P., Cecillon, S., Clark, I. M., Constancias, F., Simonet, P., & Vogel, T. M. (2011). Accessing the soil metagenome for studies of microbial diversity. *Applied and Environmental Microbiology*, *77*(4), 1315-1324..

Delmont, T. O., Prestat, E., Keegan, K. P., Faubladier, M., Robe, P., Clark, I. M., & Vogel, T. M. (2012). Structure, fluctuation and magnitude of a natural grassland soil metagenome. *The ISME journal*, *6*(9), 1677-1687.

Delmont, T. O., Simonet, P., & Vogel, T. M. (2012). Describing microbial communities and performing global comparisons in the 'omic era. *The ISME journal*, *6*(9), 1625-1628.

Dequiedt, S., Thioulouse, J., Jolivet, C., Saby, N., Lelievre, M., Maron, P. A., Ranjard, L. (2009). Biogeographical patterns of soil bacterial communities. *Environmental microbiology reports*, *1*(4), 251-255.

Edwards U, Rogall T, Blocker H, Emde M, Bottger EC (1989). Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res*. 17: 7843-7853.

Egert, M., Wagner, B., Lemke, T., Brune, A., & Friedrich, M. W. (2003). Microbial community structure in midgut and hindgut of the humus-feeding larva of Pachnoda ephippiata (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology*, *69*(11), 6659-6668.

Eilers, K. G., Debenport, S., Anderson, S., & Fierer, N. (2012). Digging deeper to find unique microbial communities: the strong effect of depth on the structure of bacterial and archaeal communities in soil. *Soil Biology and Biochemistry*, *50*, 58-65.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 677.

Fierer, N., Schimel, J. P., & Holden, P. A. (2003). Variations in microbial community composition through two soil depth profiles. *Soil Biology and Biochemistry*, *35*(1), 167-176.

Fierer, N., & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 626-631.

Fierer, N., Carney, K. M., Horner-Devine, M. C., & Megonigal, J. P. (2009). The biogeography of ammonia-oxidizing bacterial communities in soil. *Microbial ecology*, *58*(2), 435-445.

Fierer, N., Strickland, M. S., Liptzin, D., Bradford, M. A., & Cleveland, C. C. (2009). Global patterns in belowground communities. *Ecology Letters*, *12*(11), 1238-1249.

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., & Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial

communities and their functional attributes. *Proceedings of the National Academy of Sciences*, *109*(52), 21390-21395.

Filion, M., St-Arnaud, M., & Fortin, J. A. (1999). Direct interaction between the arbuscular mycorrhizal fungus Glomus intraradices and different rhizosphere microorganisms. *New phytologist*, *141*(3), 525-533.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E.,& Bateman, A. (2009). The Pfam protein families database. *Nucleic acids research*, gkp985..

Foster, J. A., Bunge, J., Gilbert, J. A., & Moore, J. H. (2012). Measuring the microbiome: perspectives on advances in DNA-based techniques for exploring microbial life. *Briefings in bioinformatics*, bbr080.

Franklin, J. (2010). *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Franklin, R. B., & Mills, A. L. (2003). Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. *FEMS Microbiology Ecology*, *44*(3), 335-346.

Franklin, R. B., & Mills, A. L. (2009). Importance of spatially structured environmental heterogeneity in controlling microbial community composition at small spatial scales in an agricultural field. *Soil Biology and Biochemistry*,*41*(9), 1833-1840.

Frey, S. D., Elliott, E. T., & Paustian, K. (1999). Bacterial and fungal abundance and biomass in conventional and no-tillage agroecosystems along two climatic gradients. *Soil Biology and Biochemistry*, *31*(4), 573-585.

Frey, S. D., Drijber, R., Smith, H., & Melillo, J. (2008). Microbial biomass, functional capacity, and community structure after 12 years of soil warming.*Soil Biology and Biochemistry*, *40*(11), 2904-2907.

Fritze, H., Pietikäinen, J., & Pennanen, T. (2000). Distribution of microbial biomass and phospholipid fatty acids in Podzol profiles under coniferous forest.*European Journal of Soil Science*, *51*(4), 565-573.

Frisli, T., Haverkamp, T. H. A., Jakobsen, K. S., Stenseth, N. C., & Rudi, K. (2013). Estimation of metagenome size and structure in an experimental soil microbiota from low coverage next-generation sequence data. *Journal of applied microbiology*, *114*(1), 141-151.

Frostegård, Å., Tunlid, A., & Bååth, E. (1993). Phospholipid fatty acid composition, biomass, and activity of microbial communities from two soil types experimentally exposed to different heavy metals. *Applied and Environmental Microbiology*, *59*(11),

3605-3617.

Gans, J., Wolinsky, M., & Dunbar, J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science*,*309*(5739), 1387-1390.

García-Salamanca, A., Molina-Henares, M. A., Dillewijn, P., Solano, J., Pizarro-Tobías, P., Roca, A., ... & Ramos, J. L. (2013). Bacterial diversity in the rhizosphere of maize and the surrounding carbonate-rich bulk soil. *Microbial biotechnology*, *6*(1), 36-46.

Ge, Y., Zhang, J. B., Zhang, L. M., Yang, M., & He, J. Z. (2008). Long-term fertilization regimes affect bacterial community structure and diversity of an agricultural soil in northern China. *Journal of Soils and Sediments*, *8*(1), 43-50.

Ghosh, A. K., Pattnaik, A. K., & Ballatore, T. J. (2006). Chilika Lagoon: Restoring ecological balance and livelihoods through re-salinization. *Lakes & Reservoirs: Research & Management*, *11*(4), 239-255.

Griffiths G., Thomson R.I., James B.C., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of British soils. *Environmental microbiology*, *13*(6), 1642-1654.

Grundmann, G. L., Dechesne, A., Bartoli, F., Flandrois, J. P., Chasse, J. L., & Kizungu, R. (2001). Spatial modeling of nitrifier microhabitats in soil. *Soil Science Society of America Journal*, *65*(6), 1709-1716.

Grundmann, G. L. (2004). Spatial scales of soil bacterial diversity–the size of a clone. *FEMS Microbiology Ecology*, *48*(2), 119-127.

Hartmann, M., Lee, S., Hallam, S. J., & Mohn, W. W. (2009). Bacterial, archaeal and eukaryal community structures throughout soil horizons of harvested and naturally disturbed forest stands. *Environmental Microbiology*,*11*(12), 3045-3062.

Hasegawa, M., Ito, M. T., Yoshida, T., Seino, T., Chung, A. Y., & Kitayama, K. (2014). The effects of reduced-impact logging practices on soil animal communities in the Deramakot Forest Reserve in Borneo. *Applied Soil Ecology*, *83*, 13-21.

Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., & Andersson, A. F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME journal*, *5*(10), 1571-1579.

Holden, P. A., Ritz, K., & Young, I. (2011). How do the microhabitats framed by soil structure impact soil bacteria and the processes that they regulate. *The architecture and biology of soils: life in inner space*, 118-149.

Hou, L. B., Yanagisawa, Y., Yachi, S., Kaneko, N., & Nakamori, T. (2014). Biomass estimation of the terrestrial ecotoxicological species Folsomia candida (Collembola)

using a real-time polymerase chain reaction.*Ecotoxicology and environmental safety*, *101*, 59-63.

Huber, J. A., Butterfield, D. A., & Baross, J. A. (2003). Bacterial diversity in a subseafloor habitat following a deep-sea volcanic eruption. *FEMS Microbiology Ecology*, *43*(3), 393-409.

Janssen, P. H. (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and environmental microbiology*,*72*(3), 1719-1728.

Jones, R. T., Robeson, M. S., Lauber, C. L., Hamady, M., Knight, R., & Fierer, N. (2009). A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses. *The ISME journal*, *3*(4), 442-453.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., & Yamanishi, Y. (2008). KEGG for linking genomes to life and the environment.*Nucleic acids research*, *36*(suppl 1), D480-D484.

Kelley, D. S., Baross, J. A., & Delaney, J. R. (2002). Volcanoes, fluids, and life at mid-ocean ridge spreading centers. *Annual Review of Earth and Planetary Sciences*, *30*(1), 385-491.

Kowalchuk, G. A., Bodelier, P. L., Heilig, G. H. J., Stephen, J. R., & Laanbroek, H. J. (1998). Community analysis of ammonia-oxidising bacteria, in relation to oxygen availability in soils and root-oxygenated sediments, using PCR, DGGE and oligonucleotide probe hybridisation. *FEMS Microbiology Ecology*, *27*(4), 339-350.

Kramer, S., Marhan, S., Haslwimmer, H., Ruess, L., & Kandeler, E. (2013). Temporal variation in surface and subsoil abundance and function of the soil microbial community in an arable soil. *Soil Biology and Biochemistry*, *61*, 76-85.

Kravchenko, A. N., Negassa, W. C., Guber, A. K., Hildebrandt, B., Marsh, T. L., & Rivers, M. L. (2014). Intra-aggregate pore structure influences phylogenetic composition of bacterial community in macroaggregates. *Soil Science Society of America Journal*, *78*(6), 1924-1939.

Krömer, T., Kessler, M., Robbert Gradstein, S., & Acebey, A. (2005). Diversity patterns of vascular epiphytes along an elevational gradient in the Andes.*Journal of Biogeography*, *32*(10), 1799-1809.

Kumar, R., & Pattnaik, A. K. (2010). Integrated planning management for Chilika Lake. *Chilika Newslett*, *5*, 1-12.

Labbé, D., Margesin, R., Schinner, F., Whyte, L. G., & Greer, C. W. (2007). Comparative phylogenetic analysis of microbial communities in pristine and hydrocarbon-contaminated Alpine soils. *FEMS microbiology ecology*, *59*(2), 466-475.

Larose, C., Berger, S., Ferrari, C., Navarro, E., Dommergue, A., Schneider, D., & Vogel, T. M. (2010). Microbial sequences retrieved from environmental samples from seasonal Arctic snow and meltwater from Svalbard, Norway.*Extremophiles*, *14*(2), 205-212.

Latham, J., Cumani, R., Rosati, I., & Bloise, M. (2014). FAO global land cover (GLC-SHARE) Beta-Release 1.0 Database. *Division LaW*.

Lauber, C. L., Ramirez, K. S., Aanderud, Z., Lennon, J., & Fierer, N. (2013). Temporal variability in soil microbial communities across land-use types. *The ISME journal*, *7*(8), 1641-1650.

Lavelle P (1997). Faunal activities and soil processes : adaptative strategies that determine ecosystem function. *Advances in ecological recearch*. 27: 93-130.

Leff, J. W., Jones, S. E., Prober, S. M., Barberán, A., Borer, E. T., Firn, J. L., ... & Fierer, N. (2015). Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academy of Sciences*, *112*(35), 10967-10972.

Maccario, L., Vogel, T. M., & Larose, C. (2014). Potential drivers of microbial community structure and function in Arctic spring snow. *Frontiers in microbiology*, *5*.

Maghsoudi, S., Vossoughi, M., Kheirolomoom, A., Tanaka, E., & Katoh, S. (2001). Biodesulfurization of hydrocarbons and diesel fuels by Rhodococcus sp. strain P32C1. *Biochemical Engineering Journal*, *8*(2), 151-156.

Mahmoudi, N., Slater, G. F., & Fulthorpe, R. R. (2011). Comparison of commercial DNA extraction kits for isolation and purification of bacterial and eukaryotic DNA from PAH-contaminated soils. *Canadian journal of microbiology*, *57*(8), 623-628.

Maul, J. E., Buyer, J. S., Lehman, R. M., Culman, S., Blackwood, C. B., Roberts, D. P., ... & Teasdale, J. R. (2014). Microbial community structure and abundance in the rhizosphere and bulk soil of a tomato cropping system that includes cover crops. *Applied Soil Ecology*, *77*, 42-50.

McCabe, D. P. (2011). Spatial location of active soil bacteria and their association with soil organic matter fractions (Doctoral dissertation, University of Reading).

Mende, D. R., Waller, A. S., Sunagawa, S., Järvelin, A. I., Chan, M. M., Arumugam, M., & Bork, P. (2012). Assessment of metagenomic assembly using simulated next generation sequencing data. *PloS one*, *7*(2), e31386.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., ... & Edwards, R. A. (2008). The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, *9*(1), 386.

Milcu, A., Allan, E., Roscher, C., Jenkins, T., Meyer, S. T., Flynn, D., ... & Eisenhauer, N. (2013). Functionally and phylogenetically diverse plant communities key to soil biota. *Ecology*, *94*(8), 1878-1885.

Mitchell, R. J., Hester, A. J., Campbell, C. D., Chapman, S. J., Cameron, C. M., Hewison, R. L., & Potts, J. M. (2012). Explaining the variation in the soil microbial community: do vegetation composition and soil chemistry explain the same or different parts of the microbial variation?. *Plant and soil*, *351*(1-2), 355-362.

Mummey, D., Holben, W., Six, J., & Stahl, P. (2006). Spatial stratification of soil bacterial populations in aggregates of diverse soils. *Microbial Ecology*,*51*(3), 404-411.

Muyzer et al., 1993, Phylogenetic relationships ofThiomicrospira species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. Volume 164, pp 165-172

Neufeld, J. D., & Mohn, W. W. (2005). Unexpectedly high bacterial diversity in arctic tundra relative to boreal forest soils, revealed by serial analysis of ribosomal sequence tags. *Applied and Environmental Microbiology*, *71*(10), 5710-5718.

Nielsen, U.N., Ayres, E. & Wall, D.H., 2011. Soil biodiversity and carbon cycling: a review and synthesis of studies examining diversity–function relationships. *European Journal of Soil Science*, *62*(1), 105-116.

O'donnell, A. G., Seasman, M., Macrae, A., Waite, I., & Davies, J. T. (2001). Plants and fertilisers as drivers of change in microbial community structure and function in soils. *Plant and Soil*, *232*(1-2), 135-145.

Ollivier, J., Töwe, S., Bannert, A., Hai, B., Kastl, E. M., Meyer, A., ... & Schloter, M. (2011). Nitrogen turnover in soil and global change. *FEMS microbiology ecology*, *78*(1), 3-16.

Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., ... & Vonstein, V. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic acids research*, *33*(17), 5691-5702.

Ovreas and Torsvik, 1998, Microbial diversity and community structure in two different agricultural soil communities *Microbiol. Ecol.*, 36, pp. 303–315

Parkin, T. B. (1987). Soil microsites as a source of denitrification variability.*Soil Science Society of America Journal*, *51*(5), 1194-1199.

Parks and Beiko, 2010, Identifying biologically relevant differences between metagenomic communities. *Bioinformatics*, 26(6), 715-721.

Paul and Clark, 1989, Soil Microbiology and Biochemistry, Academic Press, London

Peiffer, J. A., Spor, A., Koren, O., Jin, Z., Tringe, S. G., Dangl, J. L., & Ley, R. E. (2013). Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proceedings of the National Academy of Sciences*,*110*(16), 6548-6553.

Perkins, L. B., Blank, R. R., Ferguson, S. D., Johnson, D. W., Lindemann, W. C., & Rau, B. M. (2013). Quick start guide to soil methods for ecologists.*Perspectives in Plant Ecology, Evolution and Systematics*, *15*(4), 237-244.

Philippot, L., Čuhel, J., Saby, N., Chèneby, D., Chroňáková, A., Bru, D., ... & Šimek, M. (2009). Mapping field-scale spatial patterns of size and activity of the denitrifier community. *Environmental microbiology*, *11*(6), 1518-1526.

Philippot, L., Bru, D., Saby, N., Čuhel, J., Arrouays, D., Šimek, M., & Hallin, S. (2009). Spatial patterns of bacterial taxa in nature reflect ecological traits of deep branches of the 16S rRNA bacterial tree. *Environmental Microbiology*,*11*(12), 3096-3104.

Philippot, L., Andersson, S. G., Battin, T. J., Prosser, J. I., Schimel, J. P., Whitman, W. B., & Hallin, S. (2010). The ecological coherence of high bacterial taxonomic ranks. *Nature Reviews Microbiology*, *8*(7), 523-529.

Piza, F. F., Prado, P. I., & Manfio, G. P. (2005). Investigation of bacterial diversity in Brazilian tropical estuarine sediments reveals high actinobacterial diversity. *Antonie Van Leeuwenhoek*, *86*(4), 317-328.

Pramanik, A., Basak, P., Banerjee, S., Sengupta, S., Chattopadhyay, D., & Bhattacharyya, M. (2015). Pyrosequencing based profiling of the bacterial community in the Chilika lake, the largest lagoon of India. *Genomics Data*, *4*, 112-114.

Prober, S. M., Leff, J. W., Bates, S. T., Borer, E. T., Firn, J., Harpole, W. S., Fierer, N. (2015). Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. *Ecology letters*, *18*(1), 85-95.

Prosser J., Embley, T. M., Webster, G., Hails, R. S., Beringer, J. E., & Godfray, H. C. J., 2003, The influence of selection pressures on species diversity, functional gene diversity and activity of ammonia-oxidizing bacteria. In 'Genes in the Environment'. *The Fifteenth Special Symposium of the British Ecological Society.* (pp. 187-202). Blackwell Publishing.

Pushparajan, N., & Soundarapandian, P. (2010). Recent farming of marine black tiger shrimp, Penaeus Monodon (Fabricius) in South India. *African Journal of Basic & Applied Sciences*, *2*(1-2), 33-36.

Raaijmakers, J. M., Weller, D. M., & Thomashow, L. S. (1997). Frequency of antibiotic-producing Pseudomonas spp. in natural environments. *Applied and Environmental Microbiology*, *63*(3), 881-887.

Ramette, A., & Tiedje, J. M. (2007). Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem.*Proceedings of the National Academy of Sciences*, *104*(8), 2761-2766.

Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiology Ecology*, *62*(2), 142-160.

Ranjard, L., & Richaume, A. (2001). Quantitative and qualitative microscale distribution of bacteria in soil. *Research in Microbiology*, *152*(8), 707-716.

Ranjard, L., Nazaret, S., Gourbière, F., Thioulouse, J., Linet, P., & Richaume, A. (2000). A soil microscale study to reveal the heterogeneity of Hg (II) impact on indigenous bacteria by quantification of adapted phenotypes and analysis of community DNA fingerprints. *FEMS Microbiology Ecology*, *31*(2), 107-115.

Ranjard L, Poly F, Lata JC, Mougel C, Thioulouse J, Nazaret S (2001). Chracterization of bacterial and Fungal Soil Communities by Automated Ribosomal Intergenic Spacer Analysis Finderprintints : Biological and methodological Variability. *Appl Environ Microbiol*. 67: 4479-4487

Remenant, B., Grundmann, G. L., & Jocteur-Monrozier, L. (2009). From the micro-scale to the habitat: assessment of soil bacterial community structure as shown by soil structure directed sampling. *Soil Biology and Biochemistry*,*41*(1), 29-36.

Rennert, T., Händel, M., Höschen, C., Lugmeier, J., Steffens, M., & Totsche, K. U. (2014). A NanoSIMS study on the distribution of soil organic matter, iron and manganese in a nodule from a Stagnosol. *European Journal of Soil Science*, *65*(5), 684-692.

Robe, P., Nalin, R., Capellano, C., Vogel, T. M., & Simonet, P. (2003). Extraction of DNA from soil. *European Journal of Soil Biology*, *39*(4), 183-190.

Rodriguez and Konstantinidis, 2014, Estimating coverage in metagenomic data sets and why it matters, *The ISME Journal* 8, 2349-2351

Saadoun, I. (2002). Isolation and characterization of bacteria from crude petroleum oil contaminated soil and their potential to degrade diesel fuel.*Journal of basic microbiology*, *42*(6), 420-428.

Sahu, B. K., Pati, P., & Panigrahy, R. C. (2014). Environmental conditions of Chilika Lake during pre and post hydrological intervention: an overview. *Journal of Coastal Conservation*, *18*(3), 285-297.

Sanguin H, Remenant B, Dechesne A, Thioulouse J, Vogel TM, Nesme X et al (2006). Potential of a 16S rRNA-based taxonomic microarray for analyzing the rhizosphere effects of maize on Agrobacterium spp. and bacterial communities. *Appl Environ Microbiol*. 72: 4302-4312.

Satyanarayana, C. H. (1999). *Hydrographic and phytoplankton characteristics of Chilka Lake, a brackishwater lagoon, on the east coast of India* (Doctoral dissertation, Ph. D. thesis, Andhra University, Visakhapatnam, India).

Saul, D. J., Aislabie, J. M., Brown, C. E., Harris, L., & Foght, J. M. (2005). Hydrocarbon contamination changes the bacterial diversity of soil from around Scott Base, Antarctica. *FEMS Microbiology Ecology*, *53*(1), 141-155.

Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., & Ye, J. (2011). Database resources of the national center for biotechnology information. *Nucleic acids research*, *39*(suppl 1), D38-D51.

Schloss, P. D., & Handelsman, J. (2006). Toward a census of bacteria in soil.*PLoS Comput Biol*, *2*(7), e92.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., & Weber, C. F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology*,*75*(23), 7537-7541.

Schramm, A., de Beer, D., van den Heuvel, J. C., Ottengraf, S., & Amann, R. (1999). Microscale Distribution of Populations and Activities ofNitrosospira and Nitrospira spp. along a Macroscale Gradient in a Nitrifying Bioreactor: Quantification by In Situ Hybridization and the Use of Microsensors. *Applied and environmental microbiology*, *65*(8), 3690-3696.

Schütz, K., Kandeler, E., Nagel, P., Scheu, S., & Ruess, L. (2010). Functional microbial community response to nutrient pulses by artificial groundwater recharge practice in surface soils and subsoils. *FEMS microbiology ecology*,*72*(3), 445-455.

Shen, C., Xiong, J., Zhang, H., Feng, Y., Lin, X., Li, X., ... & Chu, H. (2013). Soil pH drives the spatial distribution of bacterial communities along elevation on Changbai Mountain. *Soil Biology and Biochemistry*, *57*, 204-211.

Silvertown, J., Poulton, P., Johnston, E., Edwards, G., Heard, M., & Biss, P. M. (2006). The Park Grass Experiment 1856–2006: its contribution to ecology.*Journal of Ecology*, *94*(4), 801-814.

Six, J., Bossuyt, H., Degryze, S., & Denef, K. (2004). A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. *Soil and Tillage Research*, *79*(1), 7-31.

Smalla K, Wieland G, Buchner A, Zock A, Parzy J, Kaiser S et al (2001). Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Appl Environ Microbiol*. 67: 4742-4751.

Sohlenius, B., 1979. A carbon budget for nematodes, rotifers and tardigrades in a Swedish coniferous forest soil. *Holarctic Ecology*, 2, pp.30–40.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., ... & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored

"rare biosphere". *Proceedings of the National Academy of Sciences*, *103*(32), 12115-12120.

Stone, M. M., DeForest, J. L., & Plante, A. F. (2014). Changes in extracellular enzyme activity and microbial community structure with soil depth at the Luquillo Critical Zone Observatory. *Soil Biology and Biochemistry*, *75*, 237-247.

Tang, Z., & Fang, J. (2003). A review on the elevational patterns of plant species diversity. *Biodiversity Science*, *12*(1), 20-28.

Tatusov, R. L., Natale, D. A., Garkavtsev, I. V., Tatusova, T. A., Shankavaram, U. T., Rao, B. S., ... & Koonin, E. V. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic acids research*, *29*(1), 22-28.

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., ... & De Kesel, A. (2014). Global diversity and geography of soil fungi.*Science*, *346*(6213), 1256688.

Teisserenc, R., Lucotte, M., Houel, S., & Carreau, J. (2010). Integrated transfers of terrigenous organic matter to lakes at their watershed level: A combined biomarker and GIS analysis. *Geochimica et Cosmochimica Acta*,*74*(22), 6375-6386.

Thioulouse J, Chesse D, Dolédec S, Olivier JM (1997). ADE-4: a multivariate analysis and graphical display software. *Stat Comput*. 7: 75-83.

Thomson, B. C., Tisserant, E., Plassart, P., Uroz, S., Griffiths, R. I., Hannula, S. E., ... & Lemanceau, P. (2015). Soil conditions and land use intensification effects on soil microbial communities across a range of European field sites.*Soil Biology and Biochemistry*, *88*, 403-413.

Thureborn, P., Lundin, D., Plathan, J., Poole, A. M., Sjöberg, B. M., & Sjöling, S. (2013). a metagenomics transect into the deepest point of the Baltic Sea reveals clear stratification of microbial functional capacities. *PloS one*, *8*(9).

Timonen, S., Finlay, R. D., Olsson, S., & Söderström, B. (1996). Dynamics of phosphorus translocation in intact ectomycorrhizal systems: non-destructive monitoring using a β-scanner. *FEMS microbiology ecology*, *19*(3), 171-180.

Torsvik, V., Goksøyr, J., & Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Applied and environmental microbiology*, *56*(3), 782-787.

Tringe, S. G., Von Mering, C., Kobayashi, A., Salamov, A. A., Chen, K., Chang, H. W., ... & Rubin, E. M. (2005). Comparative metagenomics of microbial communities. *Science*, *308*(5721), 554-557

Torsvik, V. & Øvreås, L. Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol*. 5, 240–5 (2002).

Turnbull, M.S., George, P.B.L. & Lindo, Z., 2014. Weighing in: Size spectra as a standard tool in soil community analyses. *Soil Biology and Biochemistry*, 68(C), pp.366–372.

Urban, N. R., Dinkel, C., & Wehrli, B. (1997). Solute transfer across the sediment surface of a eutrophic lake: I. Porewater profiles from dialysis samplers. *Aquatic Sciences*, *59*(1), 1-25.

Uroz, S., Buée, M., Murat, C., Frey-Klett, P., & Martin, F. (2010). Pyrosequencing reveals a contrasted bacterial diversity between oak rhizosphere and surrounding soil. *Environmental Microbiology Reports*, *2*(2), 281-288.

Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of bacterial diversity in soil. *FEMS microbiology reviews*, *37*(6), 936-954.

Vogel, J., Normand, P., Thioulouse, J., Nesme, X., & Grundmann, G. L. (2003). Relationship between spatial and genetic distance in Agrobacterium spp. in 1 cubic centimeter of soil. *Applied and environmental microbiology*,*69*(3), 1482-1487

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology*, *73*(16), 5261-5267.

Webster, G., O'Sullivan, L. A., Meng, Y., Williams, A. S., Sass, A. M., Watkins, A. J. & Weightman, A. J. (2015). Archaeal community diversity and abundance changes along a natural salinity gradient in estuarine sediments. *FEMS microbiology ecology*, *91*(2), 1-18.

Wessén, E., Söderström, M., Stenberg, M., Bru, D., Hellman, M., Welsh, A. & Hallin, S. (2011). Spatial distribution of ammonia-oxidizing bacteria and archaea across a 44-hectare farm related to ecosystem functioning. *The ISME journal*, *5*(7), 1213-1225.

Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences*, *95*(12), 6578-6583

Wu, T., Ayres, E., Li, G., Bardgett, R. D., Wall, D. H., & Garey, J. R. (2009). Molecular profiling of soil animal diversity in natural ecosystems: Incongruence of molecular and morphological results. *Soil Biology and Biochemistry*, *41*(4), 849-857.

Xia X, McClelland M, Wang Y (2005). WebArray: an online platform for microarray data analysis. *BMC Bioinformatics.* 6: 306.

Yergeau E, Kang S, He Z, Zhou J, Kowalchuk GA (2007). Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. *Isme Journal.* 1: 163-179.

Young, I. M., Crawford, J. W., Nunan, N., Otten, W., & Spiers, A. (2008). Microbial distribution in soils: physics and scaling. *Advances in agronomy*,*100*, 81-121.

Young and Crawford, 2004, Interactions and Self-Organization in the Soil-Microbe Complex. *Science* Vol. 304 no. 5677 pp. 1634-1637

Zaitseva, S. V., Abidueva, E. Y., Namsaraev, B. B., Wang, L., & Wu, L. (2014). Microbial community of the bottom sediments of the brackish Lake Beloe (Transbaikal region). *Microbiology*, *83*(6), 861-868.

Zhang, W., Parker, K. M., Luo, Y., Wan, S., Wallace, L. L., & Hu, S. (2005). Soil microbial responses to experimental warming and clipping in a tallgrass prairie. *Global Change Biology*, *11*(2), 266-277.

Zhang, L. M., Wang, M., Prosser, J. I., Zheng, Y. M., & He, J. Z. (2009). Altitude ammonia-oxidizing bacteria and archaea in soils of Mount Everest.*FEMS Microbiology Ecology*, *70*(2), 208-217.

Annexe

Chapter 1

Abstract

Two terms, soil and metagenome, need to be defined in this context in order to evaluate the size of the soil metagenome. Soils are highly complex environment in term of biotic and abiotic parameters. Pedologist and soil scientist have defined major group of soil, depending on physico-chemical characteristic. For microbiologist those definitions of soil are meaningless because of the extremely complex composition and distribution of the soil microorganism at different spatial scale level. But, some recent research allows us to understand some trend on microbial spatial distribution in soil. From the macro to the micro scale, soil microorganisms seem to be spatially organized. Soil is also a changing environment, metagenomic data set produced are a picture of a single moment, so with the spatial dimension we also have to consider temporal dimension. The term Metagenome is the total of all the genome present in sample. As it's not possible to access the entire genome of soil (Bias during extraction, PCR, sequencing, data analysis…), the result of a sequencing effort is a metagenomic dataset. So the size of a soil metagenome is the sum of the size of Archaea's, Prokaryotes and Eukaryotes (including macro-organisms) genomes present in our sample. For the purpose of this review we while here focus on the microorganism metagenome and more specifically the bacteria metagenom

Chapter 3

Abstract

Soils are probably the most microbial diverse ecosystem on Earth. Although considerable sequencing of DNA and rRNA from different soils, much remains to be explored in terms of how these communities are structured, the extent of their interactions and their role in ecosystem functioning. The spatial distribution of bacterial communities inhabiting the soil shows high heterogeneity, at different scale, but is still almost unexplored. Some studies have attempted to link the spatial distribution of soil microorganisms with soil physicochemical parameters (*e.g.*, relationship between soil pH v and *Acidobacter* abundance). In this project we hypothesize that heterogeneity of the bacterial community composition appears at the same scale level of the heterogeneity of soil physicochemical properties. For the first time in large scale study, a combination of phylogenetic microarray analysis, physical and chemical analysis and large scale geographic information system (GIS) analysis, have been use to study the spatial distribution of bacteria in soil in order to understand the relationship of bacterial composition of soils (from large region in the northern France) and soils factor. The multivariate analysis of phylogenetic microarray results and physical and chemical analysis didn't gave any evidence of specific soils characteristics associated with specific bacterial community structure, especially for higher taxonomic rank. In an other hand, we were able to couple taxonomic analyses of microbial community structure and geographical information systems (GIS) to demonstrate the complexity of parameters related to shifts in community structure over large distances with the example of forest versus pH effects on Acidobacteria.

Most abundant (number of hit) Phyla per sample.

Phylum /samples	Hits	P124	P126	P128	P130	P132	P136	P138	P192	P194	P196	P198	P200
Proteobacteria	13625	3068	2010	1481	2658	3191	2075	3436	1714	2101	2823	799	3011
Actinobacteria	3550	1375	715	773	1391	1753	886	1556	683	769	1521	613	1605
Firmicutes	5154	1301	1038	469	1143	1554	782	1221	860	767	1318	662	1193
root	2609	693	491	347	623	760	530	708	424	468	709	252	744
Cyanobacteria	1588	370	202	70	265	220	78	236	72	141	204	19	255
Spirochaetes	317	115	196	105	200	208	115	115	110	113	208	105	206
Bacteroidetes	1712	118	98	81	132	140	92	129	70	102	174	66	170
Euryarchaeota	361	101	15	2	33	119	101	119	2	15	33	2	144
Tenericutes	314	85	55	8	85	84	48	85	47	55	84	47	85
Crenarchaeota	92	60	60	7	57	60	7	20	20	60	90	7	90
Thermotogae	73	44	15	15	41	41	15	41	0	24	41	15	41
Planctomycetes	170	32	8	7	8	45	30	32	7	8	31	7	9
Thermodesulfobacteria	19	17	17	17	17	18	17	17	17	17	17	17	18
Acidobacteria	45	20	20	10	12	21	15	18	19	19	17	4	11
Deinococcus-Thermus	165	25	24	0	25	60	0	29	0	24	28	0	31
Verrucomicrobia	51	16	4	16	16	16	16	17	4	4	16	7	17
Chloroflexi	69	13	8	8	18	19	8	18	8	12	9	5	18
Chlorobi	28	6	6	6	6	6	6	6	6	6	6	6	6
Nitrospirae	14	4	1	1	8	4	1	4	0	1	4	1	10
Fusobacteria	67	3	1	1	2	27	1	2	1	1	3	1	3
Synergistetes	4	2	2	2	2	2	2	2	2	2	2	0	3
Deferribacteres	14	2	2	1	2	2	2	2	2	1	2	0	6
Gemmatimonadetes	5	1	1	1	3	3	3	3	0	0	3	1	3
Aquificae	12	2	1	1	1	3	1	1	1	1	1	1	2
Candidatus Poribacteria	1	1	1	1	1	1	1	1	1	1	1	0	1
Elusimicrobia	1	1	1	1	1	1	1	1	1	1	1	1	1
Fibrobacteres	6	2	0	0	2	2	0	0	0	0	2	0	2
Chlamydiae	72	0	0	0	0	0	0	1	0	0	0	0	0
Lentisphaerae	9	0	0	0	0	0	0	0	0	0	0	0	4
Dictyoglomi	7	0	0	0	0	0	0	0	0	0	0	0	0
Apicomplexa	37	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	2	0	0	0	0	0	0	0	0	0	0	0	0
Ascomycota	11	0	0	0	0	0	0	0	0	0	0	0	0
Chrysiogenetes	1	0	0	0	0	0	0	0	0	0	0	0	0
Korarchaeota	6	0	0	0	0	0	0	0	0	0	0	0	0
Platyhelminthes	1	0	0	0	0	0	0	0	0	0	0	0	0

Phylum /samples	P202	P204	P276	P278	P280	P282	P284	P286	P386	P388	P390	P392	P394
Proteobacteria	2783	2378	1787	2469	3138	2845	2758	2033	2230	2638	2199	2779	2620
Actinobacteria	905	1206	867	886	1500	1664	1324	815	1452	1298	813	1228	1271
Firmicutes	1022	1009	482	884	877	873	1175	837	595	1104	1136	1152	1110
root	583	598	419	560	689	673	684	475	527	612	504	631	597
Cyanobacteria	155	63	132	269	204	204	132	158	98	144	185	111	90
Spirochaetes	210	115	105	115	208	208	115	113	113	118	113	115	209
Bacteroidetes	113	101	76	129	170	177	133	71	84	89	91	90	87
Euryarchaeota	101	106	15	119	101	119	119	15	88	119	2	88	15
Tenericutes	56	76	10	85	56	85	85	8	1	48	83	83	56
Crenarchaeota	20	20	20	20	60	90	20	20	20	20	20	60	20
Thermotogae	3	41	39	17	42	17	41	24	39	41	39	41	39
Planctomycetes	16	31	7	31	43	20	32	8	30	31	8	31	32
Thermodesulfobacteria	17	17	0	16	18	17	17	17	17	17	17	17	17
Acidobacteria	20	18	13	12	20	12	24	20	13	20	20	18	21
Deinococcus-Thermus	24	0	0	24	30	32	4	24	0	0	0	1	0
Verrucomicrobia	4	16	13	16	16	16	17	4	16	16	16	16	16
Chloroflexi	13	9	12	13	13	15	9	9	8	15	12	9	12
Chlorobi	6	6	6	6	6	6	6	6	6	6	6	6	6
Nitrospirae	4	4	8	1	7	6	4	4	4	4	1	4	4
Fusobacteria	1	1	1	2	3	4	1	1	1	1	1	2	1
Synergistetes	2	2	2	2	3	2	2	2	2	2	2	2	2
Deferribacteres	2	2	2	2	2	2	2	2	2	2	2	2	2
Gemmatimonadetes	1	3	1	3	3	3	3	1	1	1	1	3	1
Aquificae	1	1	1	1	2	2	2	1	1	1	1	1	1
Candidatus Poribacteria	1	1	1	1	1	1	1	1	1	1	1	1	1
Elusimicrobia	1	1	1	0	1	0	1	1	1	1	1	1	1
Fibrobacteres	0	0	0	2	2	2	0	0	0	0	0	0	0
Chlamydiae	0	0	0	0	0	0	1	0	0	0	0	0	0
Lentisphaerae	0	0	0	0	0	0	0	0	0	0	0	0	0
Dictyoglomi	0	0	0	0	0	0	0	0	0	0	0	0	0
Apicomplexa	0	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	0	0	0	0	0	0	0	0	0	0	0	0	0
Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0	0
Chrysiogenetes	0	0	0	0	0	0	0	0	0	0	0	0	0
Korarchaeota	0	0	0	0	0	0	0	0	0	0	0	0	0
Platyhelminthes	0	0	0	0	0	0	0	0	0	0	0	0	0

Phylum /samples	P396	P505	P507	P515	P517	P622	P624	P626	P628	P630	P632	P634
Proteobacteria	1695	2347	1742	3497	2802	3106	811	2096	2407	3433	2687	2962
Actinobacteria	641	970	664	1781	1437	1537	576	1401	1295	1861	1264	1782
Firmicutes	712	811	797	1282	1179	1180	515	758	1108	1387	1140	1147
root	373	540	413	807	642	705	204	547	574	790	629	695
Cyanobacteria	43	158	85	227	226	196	20	190	215	365	158	267
Spirochaetes	110	113	110	208	208	211	0	108	115	208	211	208
Bacteroidetes	79	109	80	153	132	147	57	86	124	194	126	151
Euryarchaeota	88	101	2	152	101	106	2	33	101	144	119	119
Tenericutes	47	56	47	85	85	85	0	36	84	85	85	85
Crenarchaeota	7	20	20	91	60	20	19	20	20	91	60	90
Thermotogae	15	39	15	41	41	41	15	41	41	41	39	41
Planctomycetes	26	8	7	45	32	32	7	8	32	45	32	32
Thermodesulfobacteria	16	17	16	17	17	17	16	17	17	18	17	18
Acidobacteria	10	21	14	19	17	22	8	10	15	15	18	17
Deinococcus-Thermus	0	24	0	25	24	3	0	24	0	34	24	26
Verrucomicrobia	16	4	13	16	16	16	0	13	16	16	16	17
Chloroflexi	8	12	8	16	9	12	4	14	8	14	9	12
Chlorobi	6	6	6	6	6	6	6	6	6	6	6	6
Nitrospirae	1	4	1	4	4	4	1	8	4	10	4	4
Fusobacteria	1	1	1	3	2	2	1	2	3	4	3	3
Synergistetes	0	2	2	2	2	2	2	2	2	3	2	2
Deferribacteres	1	2	1	2	2	2	0	2	2	2	2	2
Gemmatimonadetes	1	1	1	3	3	1	1	1	3	3	3	3
Aquificae	1	1	1	2	2	1	1	1	1	2	1	2
Candidatus Poribacteria	1	1	1	1	1	1	0	1	1	1	1	1
Elusimicrobia	0	1	0	1	1	1	0	1	1	1	1	1
Fibrobacteres	0	0	0	2	2	2	0	0	0	2	2	2
Chlamydiae	0	0	0	0	0	0	0	0	0	0	0	1
Lentisphaerae	0	0	0	0	0	0	0	0	0	0	0	0
Dictyoglomi	0	0	0	0	0	0	0	0	0	0	0	0
Apicomplexa	0	0	0	0	0	0	0	0	0	0	0	0
Arthropoda	0	0	0	0	0	0	0	0	0	0	0	0
Ascomycota	0	0	0	0	0	0	0	0	0	0	0	0
Chrysiogenetes	0	0	0	0	0	0	0	0	0	0	0	0
Korarchaeota	0	0	0	0	0	0	0	0	0	0	0	0
Platyhelminthes	0	0	0	0	0	0	0	0	0	0	0	0

Phylum /samples	P739	P741	P743	P745	P747	P749	P83	P85	P87	Mean
Proteobacteria	1396	2785	2668	2595	3104	2975	2157	1959	2147	2417,87234
Actinobacteria	526	1695	1875	1506	1266	1497	898	1225	915	1200,85106
Firmicutes	433	1162	1617	1010	1271	1131	936	673	1502	1002,02128
root	310	605	688	602	648	653	519	447	556	566,765957
Cyanobacteria	143	199	293	240	283	198	51	44	182	170,489362
Spirochaetes	105	208	190	208	197	115	112	115	112	147,446809
Bacteroidetes	74	131	153	127	152	126	92	83	98	113,361702
Euryarchaeota	2	101	15	119	101	106	88	88	2	72,2978723
Tenericutes	0	57	88	57	84	85	48	48	83	61,8297872
Crenarchaeota	20	90	90	60	87	60	7	7	20	41
Thermotogae	15	41	41	41	41	41	17	15	39	31,8723404
Planctomycetes	7	31	20	32	20	31	30	30	31	23,1702128
Thermodesulfobacteria	0	17	18	17	18	17	17	17	17	16,3191489
Acidobacteria	14	16	13	15	19	18	13	8	19	15,9361702
Deinococcus-Thermus	0	29	28	24	26	29	0	1	24	15,5319149
Verrucomicrobia	13	16	14	16	4	16	16	16	16	13,4042553
Chloroflexi	12	8	20	8	13	9	8	8	8	11,1702128
Chlorobi	6	6	6	6	6	6	6	6	6	6
Nitrospirae	3	4	8	4	10	4	1	4	1	3,93617021
Fusobacteria	1	2	3	3	2	3	2	1	25	2,87234043
Synergistetes	2	2	2	2	2	2	2	2	2	1,9787234
Deferribacteres	1	2	2	2	2	2	2	2	2	1,89361702
Gemmatimonadetes	1	1	1	1	1	1	3	3	1	1,80851064
Aquificae	1	1	2	1	2	1	1	1	1	1,27659574
Candidatus Poribacteria	1	1	1	1	1	1	1	1	1	0,95744681
Elusimicrobia	1	1	1	1	1	1	1	1	1	0,87234043
Fibrobacteres	0	2	2	2	2	2	0	0	0	0,80851064
Chlamydiae	0	0	1	0	0	0	0	0	0	0,08510638
Lentisphaerae	0	0	0	0	0	0	0	0	0	0,08510638
Dictyoglomi	0	0	1	0	0	0	0	0	0	0,0212766
Apicomplexa	0	0	0	0	0	0	0	0	0	0
Arthropoda	0	0	0	0	0	0	0	0	0	0
Ascomycota	0	0	0	0	0	0	0	0	0	0
Chrysiogenetes	0	0	0	0	0	0	0	0	0	0
Korarchaeota	0	0	0	0	0	0	0	0	0	0
Platyhelminthes	0	0	0	0	0	0	0	0	0	0

Most abundant (number of hit) species

Species	Probes_per_species	P124	P126	P128	P130	P132	P136	P138
alpha proteobacterium KC-IT-W4	9	9	9	9	9	9	9	9
Bradyrhizobium sp. KC-EP-S3	9	9	9	9	9	9	9	9
Bradyrhizobium sp. L45	9	9	9	9	9	9	9	9
Bradyrhizobium sp. RSA104	9	9	9	9	9	9	9	9
Bradyrhizobium sp. S24543	9	9	9	9	9	9	9	9
Bradyrhizobium sp. T95	9	9	9	9	9	9	9	9
Bradyrhizobium sp. T98	9	9	9	9	9	9	9	9
Nitrobacter sp. 219	9	0	9	0	9	9	0	9
Nitrobacter sp. 263	9	0	9	0	9	9	0	9
Caulobacter sp. HGR25	8	8	8	0	8	0	8	8
Paenibacillus sp. 27-9	7	7	0	7	7	7	7	7
Bradyrhizobium sp. RRM8	10	0	10	0	0	0	0	10
Bradyrhizobium sp. RSA3	10	0	10	0	0	0	0	10
Bradyrhizobium sp. RST89	10	0	10	0	0	0	0	10
Pseudomonas sp. NSJ-16	6	6	6	6	6	6	6	6
Nitrobacter sp. BS5/19	8	0	8	0	8	8	0	8
Bradyrhizobium sp. RRD24	11	0	11	0	0	0	0	11
Bradyrhizobium sp. RSS137	11	0	11	0	0	0	0	11
Bradyrhizobium sp. RST88bis	11	0	11	0	0	0	0	11
Devosia sp. 4_C16_46	6	6	0	6	6	6	6	6
Bacteroidetes bacterium CHC2	5	5	5	5	5	5	5	5
Brevundimonas sp. AbaT-2	6	6	6	0	6	0	6	6
Bradyrhizobium sp. CCBAU 65788	6	6	6	0	6	6	0	6
Bradyrhizobium sp. Pter17	6	6	6	0	6	6	0	6
Bradyrhizobium sp. S24556	6	6	6	0	6	6	0	6
Brevundimonas sp. M_5	6	6	6	0	6	0	0	6
Caulobacter sp. ptl1	5	5	5	0	5	0	5	5
Bacteroidetes bacterium CN9_LM99	4	4	4	4	4	4	4	4
Bacteroidetes bacterium O32B_LM98	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ONB11	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ONC2	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ZHC20	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ZNB16	4	4	4	4	4	4	4	4
Bradyrhizobium sp. T93	4	4	4	4	4	4	4	4
Bradyrhizobium sp. T97	4	4	4	4	4	4	4	4
Candidatus Liberibacter americanus	4	4	4	4	4	4	4	4
Brevundimonas sp. 39(2008)	4	4	4	0	4	4	4	4
groundwater biofilm bacterium V2	4	4	4	0	4	4	4	4
groundwater planktonic bacterium X1	4	4	4	0	4	4	4	4

Species	P194	P196	P198	P200	P202	P204	P276	P278	P280	P282
alpha proteobacterium KC-IT-W4	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. KC-EP-S3	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. L45	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. RSA104	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. S24543	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. T95	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. T98	9	9	9	9	9	9	9	9	9	9
Nitrobacter sp. 219	9	0	0	9	9	0	9	0	9	9
Nitrobacter sp. 263	9	0	0	9	9	0	9	0	9	9
Caulobacter sp. HGR25	8	0	0	8	8	8	8	8	8	8
Paenibacillus sp. 27-9	0	7	7	7	7	7	7	7	7	7
Bradyrhizobium sp. RRM8	10	0	0	10	10	0	10	0	10	0
Bradyrhizobium sp. RSA3	10	0	0	10	10	0	10	0	10	0
Bradyrhizobium sp. RST89	10	0	0	10	10	0	10	0	10	0
Pseudomonas sp. NSJ-16	6	6	0	6	6	6	6	6	6	6
Nitrobacter sp. BS5/19	8	0	0	8	8	0	8	0	8	8
Bradyrhizobium sp. RRD24	11	0	0	0	11	0	11	0	11	0
Bradyrhizobium sp. RSS137	11	0	0	0	11	0	11	0	11	0
Bradyrhizobium sp. RST88bis	11	0	0	0	11	0	11	0	11	0
Devosia sp. 4_C16_46	0	6	6	6	6	6	6	6	6	6
Bacteroidetes bacterium CHC2	5	5	5	5	5	5	5	5	5	5
Brevundimonas sp. AbaT-2	6	0	0	6	6	6	6	6	6	6
Bradyrhizobium sp. CCBAU 65788	6	6	0	6	6	0	6	6	6	6
Bradyrhizobium sp. Pter17	6	6	0	6	6	0	6	6	6	6
Bradyrhizobium sp. S24556	6	6	0	6	6	0	6	6	6	6
Brevundimonas sp. M_5	6	0	0	6	6	0	6	0	6	0
Caulobacter sp. ptl1	5	0	0	5	5	5	5	5	5	5
Bacteroidetes bacterium CN9_LM99	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium O32B_LM98	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ONB11	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ONC2	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ZHC20	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ZNB16	4	4	4	4	4	4	4	4	4	4
Bradyrhizobium sp. T93	4	4	4	4	4	4	4	4	4	4
Bradyrhizobium sp. T97	4	4	4	4	4	4	4	4	4	4
Candidatus Liberibacter americanus	4	4	4	4	4	4	4	4	4	4
Brevundimonas sp. 39(2008)	4	4	0	4	4	4	4	4	4	4
groundwater biofilm bacterium V2	4	4	0	4	4	4	4	4	4	4
groundwater planktonic bacterium X1	4	4	0	4	4	4	4	4	4	4

Species	P286	P386	P388	P390	P392	P394	P396	P505	P507	P515	P517
alpha proteobacterium KC-IT-W4	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. KC-EP-S3	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. L43	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. RSA104	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. S24343	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. T95	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. T98	9	9	9	9	9	9	9	9	9	9	9
Nitrobacter sp. 219	9	9	9	9	9	9	0	9	9	9	9
Nitrobacter sp. 263	9	9	9	9	9	9	0	9	9	9	9
Caulobacter sp. HGR25	8	8	8	8	8	8	0	8	8	8	8
Paenibacillus sp. Z7-9	0	7	7	7	7	7	7	7	7	7	7
Bradyrhizobium sp. RRM8	10	10	10	10	10	10	0	10	10	10	10
Bradyrhizobium sp. RSA3	10	10	10	10	10	10	0	10	10	10	10
Bradyrhizobium sp. RST89	10	10	10	10	10	10	0	10	10	10	10
Pseudomonas sp. NSJ-16	6	6	6	6	6	6	6	6	6	6	6
Nitrobacter sp. BS3/19	8	8	8	8	8	8	0	8	8	8	8
Bradyrhizobium sp. RRD24	11	11	11	11	11	11	0	11	11	11	11
Bradyrhizobium sp. RSS137	11	11	11	11	11	11	0	11	11	11	11
Bradyrhizobium sp. RST88bis	11	11	11	11	11	11	0	11	11	11	11
Devosia sp. 4_C16_46	0	6	6	6	6	6	6	6	6	6	6
Bacteroidetes bacterium CHC2	5	5	5	5	5	5	5	5	5	5	5
Brevundimonas sp. AbaT-2	6	6	6	6	6	6	0	6	6	6	6
Bradyrhizobium sp. CCBAU 65788	6	6	6	6	6	6	0	6	6	6	6
Bradyrhizobium sp. Pter17	6	6	6	6	6	6	0	6	6	6	6
Bradyrhizobium sp. S24556	6	6	6	6	6	6	0	6	6	6	6
Brevundimonas sp. M_3	6	6	6	6	6	6	0	6	6	6	6
Caulobacter sp. pti1	5	5	5	5	5	5	0	5	5	5	5
Bacteroidetes bacterium CN9_LM99	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium O32B_LM98	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ONB11	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ONC2	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ZHC20	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ZNB16	4	4	4	4	4	4	4	4	4	4	4
Bradyrhizobium sp. T93	4	4	4	4	4	4	4	4	4	4	4
Bradyrhizobium sp. T97	4	4	4	4	4	4	4	4	4	4	4
Candidatus Liberibacter americanus	4	4	4	4	4	4	4	4	4	4	4
Brevundimonas sp. 39(2008)	4	4	4	4	4	4	4	4	4	4	4
groundwater biofilm bacterium V2	4	4	4	4	4	4	4	4	4	4	4
groundwater planktonic bacterium X1	4	4	4	4	4	4	4	4	4	4	4

Species	P626	P628	P630	P632	P634	P636	P739	P741	P743	P745	P747
alpha proteobacterium KC-IT-W4	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. KC-EP-S3	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. L43	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. RSA104	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. S24343	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. T95	9	9	9	9	9	9	9	9	9	9	9
Bradyrhizobium sp. T98	9	9	9	9	9	9	9	9	9	9	9
Nitrobacter sp. 219	9	0	9	9	0	0	9	9	9	9	9
Nitrobacter sp. 263	9	0	9	9	0	0	9	9	9	9	9
Caulobacter sp. HGR25	8	8	8	0	8	0	8	8	8	8	8
Paenibacillus sp. Z7-9	7	7	7	7	7	7	7	7	0	7	7
Bradyrhizobium sp. RRM8	0	0	10	0	0	0	10	10	0	10	10
Bradyrhizobium sp. RSA3	0	0	10	0	0	0	10	10	0	10	10
Bradyrhizobium sp. RST89	0	0	10	0	0	0	10	10	0	10	10
Pseudomonas sp. NSJ-16	6	6	6	6	6	6	6	6	6	6	6
Nitrobacter sp. BS3/19	8	0	8	8	0	0	8	8	8	8	8
Bradyrhizobium sp. RRD24	0	0	0	0	0	0	0	11	0	11	11
Bradyrhizobium sp. RSS137	0	0	0	0	0	0	0	11	0	11	11
Bradyrhizobium sp. RST88bis	0	0	0	0	0	0	0	11	0	11	11
Devosia sp. 4_C16_46	6	6	6	6	6	6	6	6	0	6	6
Bacteroidetes bacterium CHC2	5	5	5	5	5	5	5	5	5	5	5
Brevundimonas sp. AbaT-2	6	6	6	0	6	0	6	6	6	6	6
Bradyrhizobium sp. CCBAU 65788	6	0	6	6	6	0	6	6	6	6	6
Bradyrhizobium sp. Pter17	6	0	6	6	6	0	6	6	6	6	6
Bradyrhizobium sp. S24556	6	0	6	6	6	0	6	6	6	6	6
Brevundimonas sp. M_3	6	0	6	0	6	0	6	6	6	6	6
Caulobacter sp. pti1	5	5	5	0	5	0	5	5	5	5	5
Bacteroidetes bacterium CN9_LM99	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium O32B_LM98	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ONB11	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ONC2	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ZHC20	4	4	4	4	4	4	4	4	4	4	4
Bacteroidetes bacterium ZNB16	4	4	4	4	4	4	4	4	4	4	4
Bradyrhizobium sp. T93	4	4	4	4	4	4	4	4	4	4	4
Bradyrhizobium sp. T97	4	4	4	4	4	4	4	4	4	4	4
Candidatus Liberibacter americanus	4	4	4	4	4	0	4	4	4	4	4
Brevundimonas sp. 39(2008)	4	4	4	4	4	0	4	4	4	4	4
groundwater biofilm bacterium V2	4	4	4	4	4	0	4	4	4	4	4
groundwater planktonic bacterium X1	4	4	4	4	4	0	4	4	4	4	4

ID	Ti83	Ti85	Ti87	Ti124	Ti126	Ti128
fract_1	209	222	183	140	262	
fract_2	280	210	356	261	256	
fract_3	433	348	400	403	332	
fract_4	63	189	39	190	97	
fract_5	15	31	22	6	53	
carbon	10,5	10,8	24,3	11,5	25,8	
N_tot	1,13	1,18	2,49	1,19	1,63	
C/N	9,29	9,15	9,76	9,66	15,83	
calc_tot	1,8	23,4	-1	1,5	-1	
ph_water	7,7	8,2	6	7,1	4,5	
ca_ech	12,8	14,8	7,49	8,26	1,34	
mg_ech	0,27	0,39	1,14	0,58	0,34	
k_ech	0,222	0,477	0,472	0,649	0,258	
0_ech	0,119	0,033	0,068	0,035	0,036	
al_ech	0,02	0,02	0,04	0,04	4,74	
mn_ech	0	0	0,02	0,01	0,454	
fe_ech	0	0	0	-0,01	0,009	
fe_lib	0	0	0	0,81	1,62	
al_tot	4,17	3,84	4,48	3,6	4,02	
ca_tot	0,58	1,58	0,42	0,44	0,18	
fe_tot	2,14	2,13	2,22	1,68	2,4	
k_tot	1,3	1,41	1,53	1,55	1,1	
mg_tot	0,32	0,35	0,33	0,26	0,24	
0_tot	0,67	0,52	0,71	0,71	0,52	
cd_tot	0,353	0,402	0,281	0,239	0,233	
co_tot	12,5	11,4	8,79	7,21	15,6	
cr_tot	55,1	54,9	51,6	38,1	42,5	
cu_tot	10,1	15	10	10,6	13,8	
mn_tot	1040	565	631	713	1090	
mo_tot	0,465	0,628	0,627	0,409	0,644	
ni_tot	23,1	21,6	16,2	13,5	27,5	
pb_tot	25,6	22,9	31,9	19	43,9	
tl_tot	0,464	0,423	0,52	0,401	0,445	
zn_tot	61,2	58,9	64,6	47,2	99,8	
b_ext	0,338	0,444	0,239	0,232	0,527	
cd_ext	0,169	0,248	0,179	0,154	0,078	
cr_ext	0,03	0,092	0,126	0,051	0,34	
cu_ext	1,78	4,59	1,94	3,07	1,67	
ni_ext	0,957	1,1	0,778	0,652	0,759	
pb_ext	5,68	7	8,48	4,93	20,3	
zn_ext	2,14	5,49	3,38	3,26	19,6	
c_n	9,29	9,15	9,76	9,66	15,83	
s_t	97,18	101,95	95,03	104,77	29,86	
x	647000	679000	711000	487000	519000	
y	2536000	2536000	2536000	2524000	2524000	

ID	Ti132	Ti136	Ti138	Ti192	Ti194
fract_1	225	245	267	180	119
fract_2	207	207	272	222	180
fract_3	345	487	405	312	267
fract_4	151	58	50	171	308
fract_5	72	3	6	115	126
carbon	14,4	10,3	10,5	52,3	18,8
N_tot	1,47	1,12	1,16	2,03	0,89
C/N	9,8	9,2	9,05	25,76	21,12
calc_tot	73,2	-1	1,3	-1	-1
ph_water	8,1	7,6	7,6	4,2	4,2
ca_ech	16,9	13,5	15,7	0,92	0,38
mg_ech	0,39	0,66	0,83	0,29	0,1
k_ech	0,316	0,627	0,657	0,181	0,091
0_ech	0,035	0,112	0,065	0,053	0,164
al_ech	0,03	0,05	0,04	3,73	2,34
mn_ech	0	0	0	0,623	0,287
fe_ech	0	0	0	0,019	0,021
fe_lib	0	0	0	0,83	0,8
al_tot	4,04	4,7	4,97	2,03	1,8
ca_tot	3,56	0,62	0,49	0,11	0,1
fe_tot	2,23	2,43	2,71	1,18	1,02
k_tot	1,26	1,54	1,46	0,43	0,39
mg_tot	0,39	0,43	0,42	0,09	0,06
0_tot	0,56	0,67	0,62	0,29	0,18
cd_tot	0,252	0,266	0,218	0,092	0,031
co_tot	9,49	10,3	11,4	3,88	2,76
cr_tot	51,3	58,7	62,2	26,7	30,1
cu_tot	16	19,9	12,8	7,52	4,71
mn_tot	600	512	537	470	234
mo_tot	0,456	0,547	0,489	0,49	0,383
ni_tot	24	25,4	25,4	9,57	5,36
pb_tot	18,3	20,6	23,8	20,6	16,4
tl_tot	0,39	0,425	0,512	0,252	0,225
zn_tot	64	67	61,3	21,9	17,4
b_ext	0,349	0,509	0,273	0,58	0,354
cd_ext	0,208	0,149	0,112	0,039	0,015
cr_ext	0,072	0,058	0,042	0,5	0,74
cu_ext	4,19	3,62	1,84	1,12	0,823
ni_ext	0,834	1,57	1,01	0,496	0,299
pb_ext	6,77	5,09	4,5	10	6,86
zn_ext	3,55	3,5	1,71	2,26	1,2
c_n	9,8	9,2	9,05	25,76	21,12
s_t	102,56	96,75	107,83	22,15	18,99
x	615000	679000	711000	487000	519000
y	2524000	2524000	2524000	2492000	2492000

ID	Ti200	Ti202	Ti204	Ti276	Ti278
fract_1	240	56	272	194	262
fract_2	185	83	249	167	205
fract_3	506	118	395	317	251
fract_4	65	535	80	138	154
fract_5	4	208	4	184	128
carbon	10,5	17,6	8,49	24,1	16,43
N_tot	1,2	0,97	1	1,74	1,81
C/N	8,75	18,14	8,49	13,85	9,08
calc_tot	1,6	-1	1,6	-1	261
ph_water	7,7	4,8	7,9	5,8	8,1
ca_ech	15,5	1,53	15,7	10,4	17,07
mg_ech	0,42	0,14	0,71	1,09	0,23
k_ech	0,305	0,084	0,544	0,187	0,33
0_ech	0,049	0,027	0,078	0,097	0,022
al_ech	0,03	1,31	0,03	0,13	0,03
mn_ech	0	0,014	0	0,129	0
fe_ech	0	0,017	0	0,013	-1,5
fe_lib	0	0,32	0	0,74	0
al_tot	5,07	1,08	4,75	2,81	2,77
ca_tot	0,65	0,09	0,62	0,39	11,12
fe_tot	2,64	0,56	2,73	1,47	1,92
k_tot	1,49	0,49	1,62	0,88	0,8
mg_tot	0,44	0,06	0,46	0,22	0,27
0_tot	0,71	0,18	0,6	0,45	0,29
cd_tot	0,14	0,089	0,259	0,389	0,77
co_tot	10,6	1,55	11,3	6,72	6,07
cr_tot	60,8	16	62,5	37,1	36,93
cu_tot	12,9	3,14	15	6,71	13,27
mn_tot	581	44,1	580	628	640,1
mo_tot	0,618	0,279	0,643	0,344	0,56
ni_tot	27,1	3,92	26,2	17,3	20,07
pb_tot	18	14,9	21,3	22,2	18,91
tl_tot	0,484	0,381	0,451	0,321	0,298
zn_tot	63,7	17,8	62,2	54,7	76,15
b_ext	0,192	0,212	0,494	0,284	0,168
cd_ext	0,162	0,061	0,141	0,232	0,231
cr_ext	0,045	0,27	0,062	0,136	-0,1
cu_ext	2,39	0,641	2,99	0,994	1,78
ni_ext	1,52	0,362	1,66	1,58	0,37
pb_ext	4,85	9,04	5,58	8,64	3,33
zn_ext	2,02	2,92	2,12	2,97	5,34
c_n	8,75	18,14	8,49	13,85	9,08
s_t	102,35	48,01	97,89	96,51	102,93
x	615000	647000	679000	519000	551000
y	2492000	2492000	2492000	2460000	2460000

ID	Ti284	Ti286	Ti386	Ti388	Ti390	Ti392
fract_1	217	192	223	147	172	
fract_2	283	282	261	183	253	
fract_3	372	444	434	220	455	
fract_4	97	55	58	389	107	
fract_5	31	27	24	61	13	
carbon	12,2	7,91	11,3	23,5	9,26	
N_tot	1,26	0,88	1,14	1,3	0,89	
C/N	9,68	8,99	9,91	18,08	10,4	
calc_tot	4,3	-1	-1	-1	-1	
ph_water	7,6	6,8	6,8	4,4	5,7	
ca_ech	13,1	8,85	11,4	0,64	4,78	
mg_ech	1,21	1,19	0,66	0,33	1,07	
k_ech	0,53	0,351	0,346	0,216	0,58	
0_ech	0,037	0,021	0,041	0,035	0,035	
al_ech	0,03	0,03	0,05	3,01	0,25	
mn_ech	0	0,01	0,076	0,063	0,04	
fe_ech	0	0	0,008	0,018	-1,5	
fe_lib	0	0	1,03	0,74	0	
al_tot	4,53	4,09	3,96	2,35	3,86	
ca_tot	0,66	0,35	0,44	0,11	0,29	
fe_tot	2,42	2,01	2,07	1,16	1,78	
k_tot	1,49	1,42	1,4	0,9	1,33	
mg_tot	0,4	0,35	0,32	0,11	0,28	
0_tot	0,61	0,68	0,65	0,32	0,63	
cd_tot	0,181	0,245	0,259	0,064	0,219	
co_tot	11,9	9,44	9,96	2,97	7,76	
cr_tot	52,8	49,8	47,6	27,5	38,68	
cu_tot	14,4	10,9	12	4,45	12,89	
mn_tot	771	637	673	152	359,5	
mo_tot	0,663	0,537	0,426	0,577	0,48	
ni_tot	22,8	20	21,4	6,65	18	
pb_tot	23	21,7	21,7	27,4	41,8	
tl_tot	0,515	0,446	0,456	0,297	0,416	
zn_tot	70,7	50	50,9	23,8	54,02	
b_ext	0,288	0,192	0,264	0,221	0,233	
cd_ext	0,308	0,13	0,151	0,048	0,114	
cr_ext	0,103	0,048	0,04	0,56	-0,1	
cu_ext	3,03	1,7	2,55	0,733	3,14	
ni_ext	2,11	1,14	1,8	0,72	1,53	
pb_ext	9,55	3,82	4	14,1	11,84	
zn_ext	9,44	1,27	1,67	2,81	2,2	
c_n	9,68	8,99	9,91	18,08	10,4	
s_t	101,2	99,16	101,2	26,43	82,57	
x	647000	679000	519000	551000	583000	
y	2460000	2460000	2428000	2428000	2428000	

ID	TÍ305	TÍ307	TÍ315	TÍ317	TÍ622
fract_1	306	150	217	231	134
fract_2	252	310	205	307	273
fract_3	196	488	243	412	429
fract_4	67	36	261	44	143
fract_5	179	16	74	6	21
carbon	26,8	9,08	8,76	9,61	9,18
N_tot	1,64	0,91	0,94	1,09	1,03
C/N	16,34	9,98	9,32	8,82	8,91
calc_tot	-1	-1	2	1	-1
ph_water	4,6	6,4	8,1	7,3	5,4
ca_ech	1,21	6,34	14,4	12,06	7,03
mg_ech	0,85	0,63	0,4	0,56	0,61
k_ech	0,156	0,3	0,45	0,45	0,27
0_ech	0,049	0,028	0,051	0,064	0,035
al_ech	4,34	0,02	-0,02	0,02	0,1
mn_ech	0,206	0,13	0	0,01	0,09
fe_ech	0,008	-0,01	-0,01	-0,01	-1,5
fe_lib	2,02	0	0	0	0
al_tot	4,65	3,69	3,36	4,26	3,09
ca_tot	0,12	0,33	0,49	0,44	0,3
fe_tot	2,77	1,67	1,97	2,06	1,58
k_tot	0,89	1,36	1,09	1,39	1,16
mg_tot	0,21	0,26	0,26	0,31	0,21
0_tot	0,32	0,7	0,34	0,55	0,55
cd_tot	0,066	0,24	0,172	0,272	0,148
co_tot	12,8	8,73	10,8	10,04	7,9
cr_tot	57,6	43,11	43,7	49	30,68
cu_tot	7,44	9,89	9,3	10,7	4,72
mn_tot	471	619,1	603,3	696,2	568,2
mo_tot	0,482	0,43	0,46	0,51	0,256
ni_tot	16,1	17,37	17	22,4	8,04
pb_tot	30	16,08	22,3	17,5	17,6
tl_tot	0,51	0,271	0,31	0,404	0,315
zn_tot	36,2	49,43	46,2	48,8	29,61
b_ext	0,206	0,189	0,159	0,247	0,175
cd_ext	0,042	0,134	0,097	0,152	0,087
cr_ext	0,27	0,041	0,158	0,163	-0,05
cu_ext	1,35	2,428	1,712	2	1,093
ni_ext	1,24	1,343	1,305	1,767	0,934
pb_ext	9	4,807	11,954	4,711	2,917
zn_ext	1,87	1,642	2,092	1,472	1,484
c_n	16,34	9,98	9,32	8,82	8,91
s_t	33,02	88,03	101,26	95,24	104,13
x	487000	519000	647000	679000	487000
y	2396000	2396000	2396000	2396000	2364000

ID	Ti630	Ti632	Ti634	Ti636	Ti739
fract_1		243	181	331	312
fract_2		300	207	248	302
fract_3		377	267	107	142
fract_4		53	172	108	60
fract_5		27	173	206	184
carbon		9,7	9,21	19	16
N_tot		1,01	0,97	1,8	1,76
C/N		9,6	9,49	10,56	9,09
calc_tot		1	8	190	689
ph_water		7,9	8,2	8,3	8,1
ca_ech		14,98	12,26	19	14,7
mg_ech		0,35	0,13	0,31	0,24
k_ech		0,33	0,17	0,311	0,662
0_ech		0,064	0,023	0,019	0,021
al_ech		0,06	-0,02	-0,02	0,05
mn_ech		0	0	0	0
fe_ech		-0,01	-0,01	0	0
fe_lib		0	0	1,67	0,7
al_tot		4,32	2,96	4,3	1,77
ca_tot		0,55	0,67	8,86	29
fe_tot		2,01	1,47	2,35	1,22
k_tot		1,3	0,87	0,48	0,34
mg_tot		0,32	0,17	0,17	0,18
0_tot		0,56	0,3	0,1	0,06
cd_tot		0,29	0,375	2,03	0,568
co_tot		10,46	9,74	15,7	5,56
cr_tot		54,5	36,1	59,2	27,1
cu_tot		11,53	11,9	16,9	9,09
mn_tot		648,7	733,3	1320	762
mo_tot		0,43	0,41	0,653	0,328
ni_tot		23,78	16,5	31,9	14
pb_tot		22,59	22,7	29,8	21,9
tl_tot		0,514	0,305	0,449	0,245
zn_tot		54,19	47,3	116	54,8
b_ext		0,266	0,112	0,189	0,238
cd_ext		0,155	0,213	0,652	0,164
cr_ext		-0,05	0,196	0,043	0,053
cu_ext		2,984	3,245	1,42	1,23
ni_ext		2,509	1,418	0,283	0,214
pb_ext		5,833	5,613	3,99	2,59
zn_ext		1,366	2,953	2,45	2,8
c_n		9,6	9,49	10,56	9,09
s_t		97,66	120,53	93,64	112,4
x		615000	647000	679000	711000
y		2364000	2364000	2364000	2364000

ID	TÍ745	TÍ747	TÍ749
fract_1		391	157
fract_2		111	312
fract_3		37	244
fract_4		103	77
fract_5		358	210
carbon		24,48	13,3
N_tot		2,35	0,91
C/N		10,42	14,62
calc_tot		8	-1
ph_water		7,5	4,5
ca_ech		27,73	0,62
mg_ech		0,94	0,19
k_ech		0,43	0,132
0_ech			0,017
al_ech		-0,02	1,97
mn_ech		0,01	0,302
fe_ech		-1,5	0,009
fe_lib		0	0,66
al_tot		5,54	2,6
ca_tot		0,85	0,11
fe_tot		2,29	1,07
k_tot		1,23	1,08
mg_tot		0,29	0,12
0_tot		0,09	0,33
cd_tot		0,277	0,06
co_tot		8,95	8,8
cr_tot		61,3	29,2
cu_tot		12,6	4,93
mn_tot		410,8	762
mo_tot		0	0,303
ni_tot		25,3	8,82
pb_tot		39,1	26,1
tl_tot		1,502	0,477
zn_tot		83,7	26,2
b_ext		0,178	0,209
cd_ext		0,153	0,032
cr_ext		0,17	0,15
cu_ext		2,082	0,478
ni_ext		1,351	0,439
pb_ext		6,365	6,74
zn_ext		2,66	1,13
c_n		10,42	14,62
s_t		0	26,71
x		615000	647000
y		2332000	2332000

Chapter 4

Abstract

Sediment and soils are among the most microbial diverse ecosystems on the Earth. The spatial distribution of bacterial communities inhabiting the sediments is highly heterogeneous at different spatial scales, but is still mostly unexplored. Some studies have suggested links between the spatial diversity of soil microorganisms and soil physicochemical parameters (*e.g.*, relationship between soil pH and *Acidobacter* abundance). In this project, we hypothesize that heterogeneity of the bacterial community composition varies at the same scale level of the heterogeneity of sediment chemical properties. Here, we focused on the large scale (km) diversity. The large scale physical and chemical characteristics that we hypothesize influence microbial communities in lake sediment at the kilometer scale are land cover, climate, pH, and salinity. We tested this by examining the spatial distribution of bacteria and physical and chemical parameters in sediment of the second largest brackish lake in the world (Chilika Lake, India). Seventy-two samples (24 stations, 3 seasons-winter, rainy and summer) of sediments from Chilika Lake were analyzed by 16S rRNA gene pyrosequencing. Land cover analyses were performed using satellite images and a digital elevation model with geographic information system (GIS), and a large set of physico-chemical analyses (e.g., pH, turbidity, salinity, conductivity) were also performed on the water column over the sediment. After a hurricane passed near the lagoon in 2013, more samples were collected to see the impact of the tropical storm on the spatial distribution of bacteria in the sediment. The results of 16S rRNA gene analysis and physical and chemical parameters used with the spatial analysis demonstrated clear spatial relationships between physico-chemical parameters (salinity), land surfaces (drainage area, type of vegetation…) and the distribution of sediment microbial communities.

Echantillon / Sample No.	44	14	6	28	20	38	7	46	21
Taxon de P1otu_table_L2	P1M30	P1M15	P1M6	P1M21	P1M33	P1M36	P1M7	P1M32	P1M34
k__Bacteria;p__Proteobacteria	0,6028	0,61884	0,4702	0,5629	0,4561	0,7468	0,373	0,73528	0,434696
k__Bacteria;p__Chloroflexi	0,0899	0,06976	0,0939	0,0843	0,2409	0,024	0,109	0,04059	0,167632
k__Bacteria;p__Bacteroidetes	0,0142	0,00959	0,0582	0,011	0,0376	0,0568	0,169	0,00743	0,114133
k__Bacteria;p__Planctomycetes	0,0353	0,03874	0,0297	0,0598	0,0399	0,0167	0,034	0,03716	0,047166
k__Bacteria;p__Acidobacteria	0,053	0,05322	0,024	0,0449	0,0153	0,0383	0,025	0,0283	0,016580
k__Bacteria;Other	0,0256	0,02595	0,1028	0,0158	0,0401	0,0145	0,115	0,02773	0,028385
k__Bacteria;p__Firmicutes	0,0228	0,00677	0,0773	0,0078	0,013	0,0207	0,005	0,00229	0,017063
k__Bacteria;p__Gemmatimonadetes	0,0364	0,06807	0,0125	0,0414	0,0091	0,0162	0,011	0,02373	0,009444
k__Bacteria;p__Actinobacteria	0,0282	0,02463	0,0131	0,0146	0,0252	0,0371	0,013	0,01515	0,018566
k__Bacteria;p__Verrucomicrobia	0,0107	0,01166	0,0226	0,0167	0,0103	0,0109	0,03	0,01086	0,017761
k__Bacteria;p__Nitrospirae	0,0132	0,01203	0,0112	0,0362	0,0088	0,0025	0,019	0,00629	0,003112
k__Bacteria;p__WS3	0,0103	0,01279	0,0069	0,0175	0,0094	0,0023	0,008	0,01029	0,006009
k__Bacteria;p__Spirochaetes	0,0061	0,00545	0,0115	0,0168	0,0141	0,0006	0,013	0,00515	0,030746
k__Bacteria;p__WS1	0,0006	0,00113	0,0002	0,0013	0,0012	6E-05	2E-04	0,00143	0,000804
k__Bacteria;p__Cyanobacteria	0,0034	0,00075	0,0063	0,0028	0,0007	0,0014	0,004	0,00086	0,003595
k__Bacteria;p__OP3	0,0063	0,00583	0,0052	0,0059	0,0136	0,0005	0,008	0,004	0,008907
k__Bacteria;p__OP8	0,0067	0,00696	0,0054	0,0149	0,0139	0,0006	0,002	0,01458	0,013307
k__Bacteria;p__Caldithrix	0,0059	0,00508	0,0066	0,0048	0,0041	0,0005	0,006	0,00629	0,005687
k__Bacteria;p__Chlorobi	0,0018	0,00113	0,0027	0,0041	0,0032	0,0015	0,008	0,002	0,011751
k__Archaea;p__Euryarchaeota	0,002	0,00019	0,0072	0,0029	0,0088	0	0,016	0	0,0050
k__Bacteria;p__Fusobacteria	0,0006	0	0,0066	0	0,0004	0,0005	0,003	0,00086	0,000107
k__Bacteria;p__Tenericutes	7E-05	0,00019	0,0044	8E-05	0,0007	0,0002	0,002	0	0,002843
k__Bacteria;p__GN04	0,002	0,00282	0,0015	0,0041	0,0023	0,0003	0,003	0,00114	0,001234
k__Bacteria;p__AC1	0,0007	0,00113	0,002	0,0062	0,0048	0	0,003	0,00086	0,00225
k__Bacteria;p__Elusimicrobia	0,0016	0,00169	0,0021	0,0022	0,0029	0,0005	0,002	0,00029	0,00225
k__Bacteria;p__TM6	0,0024	0,00282	0,001	0,0006	0,0015	0,0009	0	0,00114	0,003541
k__Bacteria;p__Lentisphaerae	0,0005	0,00038	0,0016	0,002	0,0008	6E-05	0,002	0,00114	0,001878
k__Bacteria;p__KSB3	0,0019	0,00056	0,0011	0,0004	0,002	0	0,001	0,00172	0,00225
k__Bacteria;p__NKB19	0,0014	0,0015	0,0006	0,0008	0,0007	0,0011	6E-04	0,00086	0,002039
k__Bacteria;p__BRC1	0,0009	0,00244	0,0016	0,0018	0,0011	0,0009	3E-04	0,00086	0,001717
k__Bacteria;p__	0,0004	0,00188	0,0006	0,0017	0,0012	0,0002	3E-04	0,00143	0,000160
k__Bacteria;p__NC10	0,0002	0,00056	0	0,0006	0	0,0002	0	0	0,000321
k__Bacteria;p__GN02	0,0003	0	0,0013	0,0003	0,0005	0	0,003	0	0,001448
k__Bacteria;p__Fibrobacteres	0,0003	0,00019	0,0014	0,0008	0,0043	0,0003	1E-03	0,00029	0,000965
k__Bacteria;p__SC4	0,0004	0,00094	0,0001	0,0018	0,0002	0	4E-04	0,00172	0,00026
k__Bacteria;p__[Caldithrix]	7E-05	0	0,0007	0	0,0011	6E-05	0,001	0,00086	0,004561
k__Bacteria;p__LCP-89	0,0006	0	0,0003	0,0003	0,0016	0	4E-04	0,00086	0,001073
k__Bacteria;p__WS2	0,0003	0,00019	0,0009	0,0018	0,0003	0,0003	4E-04	0	0,000482

Echantillon / Sample No.	23	3	26	68	2	16
Taxa/sample	P1M35	P1M3	P1M20	P1M38	P1M2	P1M17
k__Bacteria;p__Proteobacteria	0,77749554	0,47222884	0,78748346	0,58844016	0,65517562	0,60628953
k__Bacteria;p__Chloroflexi	0,07991212	0,12721494	0,04140669	0,10589107	0,08437093	0,1433591
k__Bacteria;p__Bacteroidetes	0,0249897	0,05284068	0,02193231	0,04809189	0,04701728	0,01581167
k__Bacteria;p__Planctomycetes	0,01331869	0,03400874	0,01852902	0,03949611	0,02648207	0,03338018
k__Bacteria;p__Acidobacteria	0,01496636	0,01597139	0,02212138	0,04320119	0,00464598	0,01914968
k__Bacteria;Other	0,01208293	0,12983711	0,0253356	0,03015932	0,06485783	0,02231202
k__Bacteria;p__Firmicutes	0,00528628	0,02749305	0,00964265	0,00148203	0,03512358	0,03004216
k__Bacteria;p__Gemmatimonadetes	0,00899355	0,01048868	0,00945358	0,04786958	0,0038097	0,01001405
k__Bacteria;p__Actinobacteria	0,0135933	0,00882002	0,01020987	0,01052242	0,008084	0,02758257
k__Bacteria;p__Verrucomicrobia	0,0090622	0,01613031	0,01153337	0,00429789	0,00724772	0,0047435
k__Bacteria;p__Nitrospirae	0,0035013	0,01295193	0,00680658	0,01756206	0,00492474	0,01370344
k__Bacteria;p__WS3	0,0045311	0,00977354	0,00813008	0,01200445	0,00436722	0,0082572
k__Bacteria;p__Spirochaetes	0,00418783	0,00802543	0,002647	0,00459429	0,00566809	0,00773015
k__Bacteria;p__WS1	0,00027461	0,0003973	0,00037814	0,00059281	0	0,00263528
k__Bacteria;p__Cyanobacteria	0,00240286	0,00198649	0,00113443	7,4102E-05	0,00195131	0,00017569
k__Bacteria;p__OP3	0,00288343	0,01493842	0,00378143	0,00348277	0,00390262	0,00579761
k__Bacteria;p__OP8	0,00583551	0,00746921	0,00813008	0,00615043	0,0123583	0,00931131
k__Bacteria;p__Caldithrix	0,00446245	0,00651569	0,00151257	0,00674324	0,00278759	0,00685172
k__Bacteria;p__Chlorobi	0,00226555	0,00707191	0,0013235	0,00281586	0,00343802	0,00140548
k__Archaea;p__Euryarchaeota	0,00137306	0,01295193	0,00018907	0,0047425	0,00250883	0
k__Bacteria;p__Fusobacteria	0,00013731	0,00063568	0	0,00066691	0,0038097	0,00035137
k__Bacteria;p__Tenericutes	0	0,00111244	0	0	0,00139379	0,00158117
k__Bacteria;p__GN04	0,00054922	0,00301947	0,00056721	0,00163023	0,00148671	0,00298665
k__Bacteria;p__AC1	0,00082384	0,00278109	0,00189072	0,00118562	0,00195131	0,00105411
k__Bacteria;p__Elusimicrobia	0,00082384	0,00103298	0,00075629	0,00066691	0,0009292	0,00158117
k__Bacteria;p__TM6	0,00061788	0,00150973	0,00056721	0,00088922	0,00074336	0,00386507
k__Bacteria;p__Lentisphaerae	0,00034327	0,00166865	0	7,4102E-05	0,00083628	0,00140548
k__Bacteria;p__KSB3	0,00020596	7,946E-05	0,00018907	0,00074102	0,00120795	0,00263528
k__Bacteria;p__NKB19	0,00034327	0,0011919	0,00018907	0,00096332	0,00037168	0,00070274
k__Bacteria;p__BRC1	0,00034327	0,00103298	0,00056721	0,00074102	0,00037168	0,00052706
k__Bacteria;p__	0,00027461	0,0007946	0,00075629	0,00111152	0,0009292	0,00105411
k__Bacteria;p__NC10	0,00068653	0	0,00094536	0,0046684	0	0,00052706
k__Bacteria;p__GN02	0,00020596	0,00031784	0	7,4102E-05	0	0,00017569
k__Bacteria;p__Fibrobacteres	0,00027461	0,0003973	0	0,0002223	0,00065044	0,00052706
k__Bacteria;p__SC4	0,00020596	0	0,00018907	0,00059281	0,0009292	0,00175685
k__Bacteria;p__[Caldithrix]	0,00041192	0,0003973	0	0,0002223	0,00037168	0,00017569
k__Bacteria;p__LCP-89	6,8653E-05	0,0007946	0,00018907	0,00029641	0,0004646	0,00158117
k__Bacteria;p__WS2	0,00020596	0,00047676	0	0,00059281	0,00027876	0,00035137

Echantillon / Sample No.	45	18	13	15	1	12	4
Taxon de P1otu_table_L2	P1M31	P1M18	P1M14	P1M16	P1M1	P1M13	P1M4
k__Bacteria;p__Proteobacteria	0,62655187	0,52003728	0,43524724	0,50201256	0,66773836	0,6922913	0,54724241
k__Bacteria;p__Chloroflexi	0,06606775	0,10251631	0,11075892	0,13701497	0,04201774	0,08916519	0,05215904
k__Bacteria;p__Bacteroidetes	0,01253977	0,05032619	0,01150154	0,03493801	0,01973392	0,01008881	0,09341599
k__Bacteria;p__Planctomycetes	0,03649635	0,02889096	0,07833726	0,03477701	0,02305987	0,03367673	0,02800342
k__Bacteria;p__Acidobacteria	0,06731549	0,05032619	0,06248868	0,04894542	0,01274945	0,01754885	0,01090209
k__Bacteria;Other	0,01403706	0,01491146	0,01775041	0,01980357	0,07827051	0,03218472	0,08379649
k__Bacteria;p__Firmicutes	0,01759311	0,0055918	0,01014309	0,01143133	0,05388027	0,01548845	0,02052159
k__Bacteria;p__Gemmatimonadetes	0,03824318	0,02935694	0,08386162	0,02978586	0,00620843	0,01449378	0,01752886
k__Bacteria;p__Actinobacteria	0,02932185	0,00885368	0,01784097	0,01432942	0,02283814	0,01428064	0,00534416
k__Bacteria;p__Verrucomicrobia	0,02632728	0,02469711	0,01295055	0,03542103	0,02206208	0,00575488	0,01795639
k__Bacteria;p__Nitrospirae	0,0071745	0,02423113	0,05261728	0,02044759	0,00199557	0,01072824	0,01603249
k__Bacteria;p__WS3	0,01253977	0,01211556	0,020739	0,01352439	0,00388027	0,01158082	0,00876443
k__Bacteria;p__Spirochaetes	0,00555244	0,01863933	0,00760732	0,01304138	0,00310421	0,00611012	0,01090209
k__Bacteria;p__WS1	0,0004991	0,00186393	0,00190183	0,00048301	0	0,00042629	0,0006413
k__Bacteria;p__Cyanobacteria	0,00199638	0	0,00063394	0,00177105	0,00232816	0,00248668	0,00149637
k__Bacteria;p__OP3	0,00349367	0,02236719	0,01141098	0,01529544	0,00288248	0,00937833	0,01154339
k__Bacteria;p__OP8	0,00854701	0,00652377	0,01313168	0,00724521	0,00210643	0,00895204	0,00470286
k__Bacteria;p__Caldithrix	0,00368083	0,0083877	0,0054338	0,00788923	0,00243902	0,00284192	0,00534416
k__Bacteria;p__Chlorobi	0,00056148	0,01584343	0,00534323	0,01159234	0,00055432	0,00191829	0,0051304
k__Archaea;p__Euryarchaeota	0,00024955	0,00046598	0,00063394	0,00048301	0,00077605	0,00213144	0,01646003
k__Bacteria;p__Fusobacteria	0,00187161	0,00139795	0,00072451	0	0,02150776	0,00092362	0,00342027
k__Bacteria;p__Tenericutes	0,00249548	0	0,00081507	0,00048301	0,00099778	0,00021314	0,0051304
k__Bacteria;p__GN04	0,00074864	0,00465983	0,00461873	0,00241507	0,00077605	0,00149201	0,0038478
k__Bacteria;p__AC1	0,00018716	0,00698975	0,00470929	0,00096603	0,00022173	0,00191829	0,00299273
k__Bacteria;p__Elusimicrobia	0,00205877	0,00326188	0,00244521	0,00483014	0,00088692	0,00127886	0,0019239
k__Bacteria;p__TM6	0,00168445	0,00186393	0,00316972	0,00177105	0,00011086	0,00170515	0,00106883
k__Bacteria;p__Lentisphaerae	0,00031193	0,00372787	0,00081507	0,00161005	0,00110865	0,00035524	0,00171013
k__Bacteria;p__KSB3	0,00155967	0,00605778	0,00018113	0,00579617	0,00055432	0,00085258	0,00021377
k__Bacteria;p__NKB19	0,00174683	0,00186393	0,00307915	0,00273708	0,00066519	0,00042629	0,0019239
k__Bacteria;p__BRC1	0,00162206	0,00046598	0,0017207	0,00080502	0,00011086	0,00071048	0,00021377
k__Bacteria;p__	0,00081103	0,00186393	0,00090563	0,00064402	0,00011086	0,00134991	0,00042753
k__Bacteria;p__NC10	0	0,00093197	0,00090563	0,00273708	0	7,1048E-05	0
k__Bacteria;p__GN02	0,00018716	0,00093197	0,00190183	0,00144904	0,00011086	0,00028419	0,00106883
k__Bacteria;p__Fibrobacteres	0,00043671	0,00046598	0,00153958	0,00177105	0,00133038	0,0001421	0,00021377
k__Bacteria;p__SC4	0,00118535	0,00186393	0,0017207	0,00096603	0	0,00035524	0,0006413
k__Bacteria;p__[Caldithrix]	6,2387E-05	0,00372787	9,0563E-05	0,00161005	0,00033259	0	0,00171013
k__Bacteria;p__LCP-89	0,00043671	0,0027959	0	0,00096603	0,00022173	0,00071048	0,00085507
k__Bacteria;p__WS2	0,00031193	0,00139795	0,00063394	0,00048301	0	0,0001421	0

Echantillon / Sample No.	25	10	9	62	5	42
Taxon de P1otu_table_L2	P1M19	P1M11	P1M10	P1M37	P1M5	P1M28
k__Bacteria;p__Proteobacteria	0,42536819	0,40911993	0,94194218	0,51399783	0,53091398	0,59337945
k__Bacteria;p__Chloroflexi	0,12030448	0,0628363	0,00929114	0,11276698	0,12920027	0,05434783
k__Bacteria;p__Bacteroidetes	0,01629985	0,0880985	0,00980993	0,07879812	0,01663306	0,04298419
k__Bacteria;p__Planctomycetes	0,07173589	0,05398997	0,0017922	0,04669965	0,04485887	0,04644269
k__Bacteria;p__Acidobacteria	0,05742181	0,01924305	0,00183936	0,02624593	0,03293011	0,02865613
k__Bacteria;Other	0,01936124	0,10943912	0,01400745	0,01508387	0,12483199	0,00889328
k__Bacteria;p__Firmicutes	0,01282476	0,02982216	0,00499929	0,06341257	0,00974462	0,05583004
k__Bacteria;p__Gemmatimonadetes	0,03003475	0,02124943	0,00183936	0,01363582	0,01276882	0,02420949
k__Bacteria;p__Actinobacteria	0,01894754	0,01121751	0,00372589	0,01816098	0,01780914	0,01828063
k__Bacteria;p__Verrucomicrobia	0,02225716	0,03666211	0,00042447	0,03831302	0,00520833	0,04001976
k__Bacteria;p__Nitrospirae	0,06106239	0,01650707	0,00056596	0,00259443	0,01478495	0,01383399
k__Bacteria;p__WS3	0,02837994	0,01541268	0,0012734	0,00699891	0,00991263	0,01432806
k__Bacteria;p__Spirochaetes	0,01869932	0,01933425	0,00146206	0,0076626	0,00386425	0,00543478
k__Bacteria;p__WS1	0,00239947	0,00118559	0	0,00024134	0,00033602	0,00049407
k__Bacteria;p__Cyanobacteria	0,00215125	0,01149111	0,00066028	0,01858332	0,00084005	0,0034585
k__Bacteria;p__OP3	0,01050803	0,01504788	0,00042447	0,00168939	0,00403226	0,00988142
k__Bacteria;p__OP8	0,02258812	0,00501596	0,00056596	0,00156872	0,00571237	0,00296443
k__Bacteria;p__Calditrhix	0,00695019	0,01705426	0,00108475	0,00681791	0,00672043	0,00839921
k__Bacteria;p__Chlorobi	0,00661923	0,00802554	0,00061312	0,00488717	0,00168011	0,00197628
k__Archaea;p__Euryarchaeota	0,00140659	0,01322389	0,00146206	0,00054302	0,00537634	0,00642292
k__Bacteria;p__Fusobacteria	0	0,00492476	0	0,00917099	0	0,00197628
k__Bacteria;p__Tenericutes	8,274E-05	0,00109439	0	0,00024134	0	0,00098814
k__Bacteria;p__GN04	0,00397154	0,00328317	0,00033014	0,00120671	0,00420027	0,00098814
k__Bacteria;p__AC1	0,00695019	0,00218878	0,00051879	0,00024134	0,00235215	0,00148221
k__Bacteria;p__Elusimicrobia	0,00289591	0,00145919	9,4326E-05	0,00078436	0,00168011	0,00592885
k__Bacteria;p__TM6	0,00066192	0,00109439	0,00014149	0,00120671	0,00050403	0,00049407
k__Bacteria;p__Lentisphaerae	0,00273043	0,00483356	4,7163E-05	0,00066369	0,00033602	0,00148221
k__Bacteria;p__KSB3	0,00124111	0,0001824	4,7163E-05	0,00072403	0,00117608	0
k__Bacteria;p__NKB19	0,00182029	0,00127679	4,7163E-05	0,00193073	0,00100806	0,00148221
k__Bacteria;p__BRC1	0,00173755	0,00136799	0	0,00060335	0,00050403	0,00148221
k__Bacteria;p__	0,00281317	0,00063839	4,7163E-05	0,00150839	0,00201613	0
k__Bacteria;p__NC10	0,00182029	0	0	0	0,00016801	0
k__Bacteria;p__GN02	0,00024822	0,00310078	0	0,00018101	0,00050403	0
k__Bacteria;p__Fibrobacteres	0,00264769	0,00264478	4,7163E-05	0	0	0,00049407
k__Bacteria;p__SC4	0,00091014	0,00072959	4,7163E-05	0,00012067	0,00117608	0,00049407
k__Bacteria;p__[Calditrhix]	0,00057918	0,00072959	0	0,00030168	0,00033602	0
k__Bacteria;p__LCP-89	0,00148933	0,00072959	4,7163E-05	0,00024134	0,00084005	0
k__Bacteria;p__WS2	0,00165481	0,000456	4,7163E-05	0,00018101	0,00067204	0,00049407

Echantillon / Sample No.	40	29	8	30	33	36	43
Taxon de P1otu_table_L2	P1M27	P1M22	P1M8	P1M23	P1M24	P1M26	P1M29
k__Bacteria;p__Proteobacteria	0,61250429	0,62805955	0,44465316	0,46517857	0,55660865	0,60840108	0,61735565
k__Bacteria;p__Chloroflexi	0,03469598	0,07784507	0,10854689	0,17732143	0,04790626	0,02777778	0,09395285
k__Bacteria;p__Bacteroidetes	0,03435246	0,02283624	0,04910774	0,01375	0,03549888	0,03884372	0,00614964
k__Bacteria;p__Planctomycetes	0,02919959	0,03570527	0,03649537	0,0425	0,03514389	0,04403794	0,06525453
k__Bacteria;p__Acidobacteria	0,02748196	0,03936412	0,01932108	0,05910714	0,05233902	0,05916893	0,03245644
k__Bacteria;Other	0,01408451	0,01703255	0,01019724	0,02321429	0,01895571	0,01151762	0,0211821
k__Bacteria;p__Firmicutes	0,10271384	0,00971486	0,09258017	0,02232143	0,02136826	0,00406504	0,00444141
k__Bacteria;p__Gemmatimonadetes	0,02988664	0,04024729	0,0119415	0,03357143	0,08030329	0,07791328	0,02938162
k__Bacteria;p__Actinobacteria	0,01236688	0,01690638	0,03005501	0,04714286	0,02291918	0,02009937	0,02528186
k__Bacteria;p__Verrucomicrobia	0,01545861	0,0122382	0,01502751	0,01178571	0,0105118	0,02122855	0,01195764
k__Bacteria;p__Nitrospirae	0,00858811	0,02977542	0,00563531	0,02910714	0,02412545	0,02168022	0,01810728
k__Bacteria;p__WS3	0,01133631	0,01009336	0,00563531	0,01232143	0,01792176	0,01738934	0,01298258
k__Bacteria;p__Spirochaetes	0,00687049	0,00908403	0,01220985	0,00857143	0,00430812	0,00203252	0,00854117
k__Bacteria;p__WS1	0,0013741	0,000757	0	0,00089286	0,00051697	0	0,00068329
k__Bacteria;p__Cyanobacteria	0,00103057	0,00050467	0,11458473	0,00178571	0,0006893	0,00112918	0,00204988
k__Bacteria;p__OP3	0,00583992	0,00555135	0,00295183	0,00553571	0,0105118	0,00903342	0,00751623
k__Bacteria;p__OP8	0,00377877	0,01286904	0,00348853	0,01321429	0,00603136	0,00383921	0,01059105
k__Bacteria;p__Caldithrix	0,00343525	0,00353268	0,00684288	0,00285714	0,00482509	0,00248419	0,00649129
k__Bacteria;p__Chlorobi	0,00343525	0,00277567	0,01113646	0,00428571	0,00224022	0,00180668	0,00034165
k__Archaea;p__Euryarchaeota	0	0,00214484	0,00187844	0,00178571	0,00861623	0,00406504	0,00136659
k__Bacteria;p__Fusobacteria	0,00961869	0	0,00201261	0	0,00034465	0,00045167	0,00034165
k__Bacteria;p__Tenericutes	0,01374098	0,0003785	0,00026835	0,00214286	0	0	0
k__Bacteria;p__GN04	0,00103057	0,00252334	0,00174426	0,0025	0,0013786	0,00090334	0,00170823
k__Bacteria;p__AC1	0	0,00277567	0,00093922	0,00267857	0,0006893	0,00112918	0,00034165
k__Bacteria;p__Elusimicrobia	0,00171762	0,00201867	0,00067087	0,00125	0,0013786	0,00112918	0,00034165
k__Bacteria;p__TM6	0,00343525	0,00138784	0,00080504	0,00196429	0,00103395	0,00112918	0,00204988
k__Bacteria;p__Lentisphaerae	0,00309172	0,00012617	0,00147592	0,00053571	0,00086162	0,00022584	0,00068329
k__Bacteria;p__KSB3	0,0013741	0,00063084	0,00093922	0,00035714	0,00017232	0	0,00170823
k__Bacteria;p__NKB19	0,00034352	0,000757	0,00161009	0,00053571	0,00120627	0,00090334	0,00307482
k__Bacteria;p__BRC1	0	0,00050467	0,00093922	0,00125	0,00086162	0,00090334	0,00170823
k__Bacteria;p__	0,00034352	0,00290184	0	0,00142857	0,00034465	0,00067751	0,00068329
k__Bacteria;p__NC10	0,00103057	0,000757	0	0,00017857	0,0013786	0,0101626	0
k__Bacteria;p__GN02	0,00034352	0,00012617	0,00174426	0,00035714	0,00017232	0,00022584	0
k__Bacteria;p__Fibrobacteres	0,00171762	0,00025233	0,00013417	0,00035714	0,00017232	0	0,00034165
k__Bacteria;p__SC4	0	0,00050467	0,00013417	0,00178571	0,00051697	0,00045167	0,00204988
k__Bacteria;p__[Caldithrix]	0,00034352	0	0,00026835	0	0,00017232	0	0,00034165
k__Bacteria;p__LCP-89	0	0,00025233	0,00120757	0,00089286	0,00017232	0	0,00136659
k__Bacteria;p__WS2	0	0	0,00026835	0,00160714	0,0006893	0,00045167	0,00034165

Echantillon / Sample No.	34	27	51	32	24	17	41
Taxon de P1otu_table_L2	P1M25	P2M28	P2M37	P2M30	P2M27	P2M24	P2M34
k__Bacteria;p__Proteobacteria	0,5	0,50217213	0,6274871	0,5090413	0,52281587	0,4791821	0,53556677
k__Bacteria;p__Chloroflexi	0	0,19149935	0,07627119	0,14631928	0,18848888	0,1957327	0,09619091
k__Bacteria;p__Bacteroidetes	0	0,01585065	0,05637436	0,02642842	0,02973324	0,04593273	0,07911886
k__Bacteria;p__Planctomycetes	0	0,04215099	0,04421518	0,0545153	0,03905832	0,03837606	0,04662689
k__Bacteria;p__Acidobacteria	0	0,03545849	0,01768607	0,05638776	0,03829397	0,03970959	0,02450665
k__Bacteria;Other	0	0,02007749	0,01731761	0,01711962	0,02086677	0,01659505	0,01890776
k__Bacteria;p__Firmicutes	0	0,0115064	0,01584377	0,00722234	0,01337614	0,01778041	0,02294631
k__Bacteria;p__Gemmatimonadetes	0	0,01115416	0,02026529	0,04402953	0,01131239	0,01170544	0,01110601
k__Bacteria;p__Actinobacteria	0	0,02900082	0,03131909	0,02904986	0,01994955	0,0188176	0,01955025
k__Bacteria;p__Verrucomicrobia	0	0,01678995	0,00773766	0,01476568	0,01521058	0,01822492	0,02872877
k__Bacteria;p__Nitrospirae	0	0,03804156	0,0110538	0,02578643	0,01238248	0,02370722	0,01046352
k__Bacteria;p__WS3	0	0,01127157	0,00478998	0,01000428	0,00978369	0,01007557	0,01037173
k__Bacteria;p__Spirochaetes	0	0,00692732	0,0092115	0,00417291	0,01146526	0,01822492	0,00614961
k__Bacteria;p__WS1	0,5	0,00011741	0,00036846	0,00053499	0,00038218	0,00029634	0,00027536
k__Bacteria;p__Cyanobacteria	0	0,00117412	0,00147384	0,00042799	0,00206375	0,00133353	0,00220285
k__Bacteria;p__OP3	0	0,01432429	0,00478998	0,0092018	0,01054804	0,00859387	0,00504819
k__Bacteria;p__OP8	0	0,01620289	0,01031688	0,0045474	0,01184744	0,00414876	0,00633318
k__Bacteria;p__Caldithrix	0	0,00457908	0,00147384	0,00567088	0,00687916	0,00651948	0,013676
k__Bacteria;p__Chlorobi	0	0,00528355	0,0018423	0,00331693	0,00726133	0,01244629	0,00780174
k__Archaea;p__Euryarchaeota	0	0,00363978	0,00073692	0,00791783	0,00244592	0,00311157	0,00091785
k__Bacteria;p__Fusobacteria	0	0	0,00073692	0,000107	0,00160514	0	0,02303809
k__Bacteria;p__Tenericutes	0	0,00011741	0,01731761	0	0,00022931	0,00029634	0,00201927
k__Bacteria;p__GN04	0	0,00082189	0,00331614	0,00224695	0,00122296	0,00237072	0,0018357
k__Bacteria;p__AC1	0	0,00551838	0,00589536	0,00128397	0,00191088	0,00192621	0,00284534
k__Bacteria;p__Elusimicrobia	0	0,00234824	0,00110538	0,00128397	0,00168157	0,00340791	0,00256999
k__Bacteria;p__TM6	0	0,00176118	0,00147384	0,00149797	0,00160514	0,00074085	0,00275356
k__Bacteria;p__Lentisphaerae	0	0,00035224	0	0,00058849	0,0012994	0,00074085	0,00027536
k__Bacteria;p__KSB3	0	0,00011741	0,00073692	0,00106998	0,00114653	0,00577863	0,00174392
k__Bacteria;p__NKB19	0	0,00046965	0,00073692	0,00192596	0,00061148	0,00074085	0,00211106
k__Bacteria;p__BRC1	0	0,00129153	0	0,00106998	0,00168157	0,00251889	0,00119321
k__Bacteria;p__	0	0,00023482	0,00221076	0,00037449	0,00061148	0,00029634	0,00036714
k__Bacteria;p__NC10	0	0,00023482	0	0,00133747	0,00091722	0,00103719	0
k__Bacteria;p__GN02	0	0,00046965	0,00073692	0,00101648	0,00045861	0	0,00100964
k__Bacteria;p__Fibrobacteres	0	0,00046965	0	0,00058849	0,00061148	0,00088902	0,00073428
k__Bacteria;p__SC4	0	0,00046965	0,00110538	0	0,00053505	0	9,1785E-05
k__Bacteria;p__[Caldithrix]	0	0	0	0,00032099	0,00145227	0,00118536	0,00073428
k__Bacteria;p__LCP-89	0	0,00129153	0	0,00026749	0,00099366	0,00059268	0,00036714
k__Bacteria;p__WS2	0	0,00035224	0,00036846	0,00080248	0,00053505	0,00029634	0,00045893

Echantillon / Sample No.	19	71	56	31	11	50	49
Taxon de P1otu_table_L2	P2M25	P2M22	P2M8	P2M29	P2M23	P2M3	P2M2
k__Bacteria;p__Proteobacteria	0,41135371	0,61228784	0,55412553	0,52792744	0,65354843	0,50883534	0,71798064
k__Bacteria;p__Chloroflexi	0,14046579	0,02931687	0,06681946	0,1027075	0,04724011	0,15060241	0,02378976
k__Bacteria;p__Bacteroidetes	0,03799127	0,05828307	0,04125529	0,13435177	0,05049555	0,02048193	0,02918396
k__Bacteria;p__Planctomycetes	0,07409025	0,04916538	0,03931594	0,0298711	0,02148593	0,05609103	0,0406639
k__Bacteria;p__Acidobacteria	0,01164483	0,03022864	0,03737659	0,02509718	0,0068726	0,02891566	0,05062241
k__Bacteria;Other	0,03100437	0,00925796	0,0190409	0,00982064	0,012226	0,03547523	0,01742739
k__Bacteria;p__Firmicutes	0,02139738	0,06571749	0,08815233	0,03389484	0,08645012	0,02637216	0,00885201
k__Bacteria;p__Gemmatimonadetes	0,00829694	0,0163417	0,02362482	0,00886585	0,00600449	0,01740295	0,01341632
k__Bacteria;p__Actinobacteria	0,01906841	0,06305232	0,03120592	0,01022983	0,01345583	0,02543507	0,03679115
k__Bacteria;p__Verrucomicrobia	0,01513828	0,0136064	0,01763047	0,04869399	0,01309412	0,00776439	0,01715076
k__Bacteria;p__Nitrospirae	0,00829694	0,00792538	0,01798307	0,00204597	0,00188092	0,01539491	0,01092669
k__Bacteria;p__WS3	0,01615721	0,00364708	0,00599436	0,00306895	0,0021703	0,01124498	0,00691563
k__Bacteria;p__Spirochaetes	0,02882096	0,00126245	0,00546544	0,01118461	0,01200897	0,01258367	0,00152144
k__Bacteria;p__WS1	0,00160116	0,00021041	0,00035261	0,0002728	0	0,00040161	0
k__Bacteria;p__Cyanobacteria	0,00334789	0,01311544	0,01339915	0,00491032	0,04846994	0,00321285	0,00705394
k__Bacteria;p__OP3	0,0279476	0,00189367	0,00440762	0,00497852	0,0005064	0,00629183	0,00124481
k__Bacteria;p__OP8	0,02663755	0,00028054	0,00405501	0,00279615	0,00115749	0,03145917	0,00055325
k__Bacteria;p__Caldithrix	0,01193595	0,00070136	0,00511283	0,00061379	0,00188092	0,00200803	0,00304288
k__Bacteria;p__Chlorobi	0,00465793	0,00070136	0,00123413	0,00879765	0,00347247	0,00267738	0,00193638
k__Archaea;p__Euryarchaeota	0,02052402	0,00154299	0	0,00109118	0,00021703	0,00053548	0,00013831
k__Bacteria;p__Fusobacteria	0,00378457	0,00596157	0,00669958	0,00279615	0,00245967	0,00040161	0,00069156
k__Bacteria;p__Tenericutes	0,00189229	0,00771497	0,00141044	0,00231876	0,00737901	0,00107095	0,00041494
k__Bacteria;p__GN04	0,00465793	0,00014027	0,00193935	0,0002728	0,00065109	0,0042838	0,00069156
k__Bacteria;p__AC1	0,00349345	0	0,0001763	0,00034099	0,00021703	0,00307898	0
k__Bacteria;p__Elusimicrobia	0,00538574	7,0136E-05	0,00123413	0,00300075	0,00028937	0,00053548	0,00027663
k__Bacteria;p__TM6	0,00262009	0,0009819	0,00088152	0,00177317	0,0005064	0,00200803	0,00096819
k__Bacteria;p__Lentisphaerae	0,00349345	0,00070136	0,00193935	0,00170497	0,00094046	0,00147256	0,00207469
k__Bacteria;p__KSB3	0,00960699	7,0136E-05	0,0001763	0,00218236	0,00014469	0,00026774	0,00082988
k__Bacteria;p__NKB19	0,00014556	0,00042082	0,00123413	0,00109118	0,00014469	0,00107095	0,00152144
k__Bacteria;p__BRC1	0,00480349	0,00014027	0,00035261	0,00184137	0,0005064	0,00093708	0,00013831
k__Bacteria;p__	0,00363901	0	0,00035261	0	0	0,00093708	0,00027663
k__Bacteria;p__NC10	0,00014556	0	0	0,00122758	0	0	0
k__Bacteria;p__GN02	0,00116448	7,0136E-05	0,00070522	0,00129578	0,00115749	0,00026774	0,00055325
k__Bacteria;p__Fibrobacteres	0,00276565	0	0,0001763	0,00075019	0,00014469	0,00013387	0
k__Bacteria;p__SC4	0,00232897	0	0,00088152	0	0	0,00200803	0
k__Bacteria;p__[Caldithrix]	0,00640466	0	0	0,00040919	0	0,00013387	0,00027663
k__Bacteria;p__LCP-89	0,00218341	0	0,00070522	0,0001364	7,2343E-05	0,00227577	0
k__Bacteria;p__WS2	0,00087336	0,00035068	0,00070522	0,00054559	0	0,00107095	0,00027663

Echantillon / Sample No.	53	66	63	58	65	47	61
Taxon de P1otu_table_L2	P2M5	P2M18	P2M15	P2M11	P2M17	P2M35	P2M14
k__Bacteria;p__Proteobacteria	0,54773869	0,47404811	0,51973311	0,72227731	0,54247643	0,58725383	0,55247247
k__Bacteria;p__Chloroflexi	0,11105528	0,05988899	0,10746735	0,02659978	0,09909563	0,10667396	0,12135608
k__Bacteria;p__Bacteroidetes	0,02462312	0,10351543	0,03478137	0,01248323	0,07845873	0,04745624	0,04567048
k__Bacteria;p__Planctomycetes	0,02663317	0,01947609	0,05153322	0,03750802	0,02482201	0,04772976	0,05463183
k__Bacteria;p__Acidobacteria	0,04070352	0,00516116	0,02938671	0,06364114	0,02193573	0,03596827	0,02267329
k__Bacteria;Other	0,04271357	0,02512416	0,02285633	0,01930817	0,01308447	0,01900985	0,01759879
k__Bacteria;p__Firmicutes	0,04773869	0,15162138	0,0330778	0,01335822	0,08399076	0,01326586	0,00896135
k__Bacteria;p__Gemmatimonadetes	0,02110553	0,00262927	0,02200454	0,01609987	0,00957283	0,02311269	0,01435975
k__Bacteria;p__Actinobacteria	0,05577889	0,01207518	0,0254117	0,04491629	0,01385415	0,02174508	0,02019002
k__Bacteria;p__Verrucomicrobia	0,01256281	0,0119778	0,02853492	0,00804993	0,03684818	0,0124453	0,03012308
k__Bacteria;p__Nitrospirae	0,00603015	0,00486902	0,00922771	0,0083416	0,00052915	0,0106674	0,00723386
k__Bacteria;p__WS3	0,01809045	0,00214237	0,01646792	0,00571662	0,00846642	0,00738512	0,0107968
k__Bacteria;p__Spirochaetes	0,00251256	0,00759568	0,01149915	0,00122499	0,00697518	0,00574398	0,01112071
k__Bacteria;p__WS1	0	0	0,00127768	0,000175	0,00024052	0,00027352	0,00064781
k__Bacteria;p__Cyanobacteria	0,00603015	0,06154445	0,01504827	0,0083416	0,01289205	0,00588074	0,01155258
k__Bacteria;p__OP3	0,00301508	0,00107119	0,00837592	0,00087499	0,00110641	0,00929978	0,00550637
k__Bacteria;p__OP8	0,00452261	0,00253189	0,0100795	0,00110832	0,00043294	0,00437637	0,00680199
k__Bacteria;p__Caldithrix	0,00854271	0,00214237	0,01263487	0,00174999	0,00582067	0,00738512	0,0131721
k__Bacteria;p__Chlorobi	0,00502513	0,01704158	0,00638842	0,00069999	0,00995767	0,00492341	0,00637011
k__Archaea;p__Euryarchaeota	0	0	0,00014196	0,000175	0,00014431	0,00232495	0,00097171
k__Bacteria;p__Fusobacteria	0,00150754	0,00262927	0,00113572	0,00046666	0,01601886	0,00013676	0,0080976
k__Bacteria;p__Tenericutes	0	0,02639011	0,00014196	0,00011667	0,0031268	0,00027352	0,00010797
k__Bacteria;p__GN04	0,00100503	0,00165547	0,00269733	0,00104999	0,00187608	0,00259847	0,00313107
k__Bacteria;p__AC1	0,00251256	9,738E-05	0,00468484	0,000175	0	0,0017779	0,00226733
k__Bacteria;p__Elusimicrobia	0,00050251	0,00019476	0,00127768	0,00023333	0,00028863	0,00205142	0,00194342
k__Bacteria;p__TM6	0,00100503	0,00038952	0,00170358	0,000175	0,00072157	0,00300875	0,00064781
k__Bacteria;p__Lentisphaerae	0	0,00019476	0,00056786	0,00029166	0,00129883	0,00136761	0,00388685
k__Bacteria;p__KSB3	0,00100503	0,0004869	0,00383305	0,00046666	0,00014431	0,00068381	0,00464263
k__Bacteria;p__NKB19	0,00150754	0,00029214	0,00127768	0,00040833	0,0010102	0,00068381	0,00107968
k__Bacteria;p__BRC1	0,00301508	0,00019476	0,00113572	0,00023333	0,00062536	0,00150438	0,00140358
k__Bacteria;p__	0	0	0,00298126	0,00011667	0,00014431	0,00123085	0,00140358
k__Bacteria;p__NC10	0	0	0	0	0	0	0
k__Bacteria;p__GN02	0	9,738E-05	0,00070982	0,00011667	0,00067347	0,00164114	0,00118765
k__Bacteria;p__Fibrobacteres	0	0,00058428	0,00141965	5,8333E-05	0,00033673	0,00027352	0,00043187
k__Bacteria;p__SC4	0	9,738E-05	0,00127768	0,00011667	0,00014431	0,00041028	0,00043187
k__Bacteria;p__[Caldithrix]	0	0,00019476	0,00070982	0,00011667	9,6209E-05	0	0,00161952
k__Bacteria;p__LCP-89	0,00050251	9,738E-05	0,00127768	0,000175	4,8105E-05	0,00027352	0,00021594
k__Bacteria;p__WS2	0,00050251	0	0,00184554	5,8333E-05	4,8105E-05	0,00082057	0,00097171

Echantillon / Sample No.	22	35	60	69	48	67	55
Taxon de P1otu_table_L2	P2M26	P2M31	P2M13	P2M20	P2M1	P2M19	P2M7
k__Bacteria;p__Proteobacteria	0,50614284	0,54423196	0,57832522	0,5702171	0,79194334	0,48303982	0,63679097
k__Bacteria;p__Chloroflexi	0,12461185	0,10968538	0,09178504	0,06936906	0,03202007	0,07972586	0,05954246
k__Bacteria;p__Bacteroidetes	0,01471581	0,06761258	0,05195192	0,02323609	0,02685554	0,01795784	0,11595111
k__Bacteria;p__Planctomycetes	0,03672202	0,04515731	0,03446906	0,04613297	0,01977276	0,05083716	0,0319649
k__Bacteria;p__Acidobacteria	0,03496692	0,02899445	0,0383431	0,0442673	0,01431312	0,04736705	0,02005641
k__Bacteria;Other	0,01930606	0,01690315	0,02552896	0,02204885	0,01106684	0,03574217	0,00752115
k__Bacteria;p__Firmicutes	0,03523694	0,00629241	0,03238303	0,02696744	0,01209975	0,00355687	0,00376058
k__Bacteria;p__Gemmatimonadetes	0,02754151	0,0127082	0,01410549	0,05393487	0,01460823	0,07755704	0,01284864
k__Bacteria;p__Actinobacteria	0,02146618	0,0159161	0,03307837	0,00729308	0,00590232	0,0090223	0,02601065
k__Bacteria;p__Verrucomicrobia	0,01134062	0,03047502	0,01490017	0,00983718	0,01254242	0,01223215	0,01316202
k__Bacteria;p__Nitrospirae	0,04644255	0,0070327	0,01281415	0,04002714	0,00604987	0,06090049	0,02256346
k__Bacteria;p__WS3	0,02119617	0,00690932	0,00824476	0,01272049	0,00693522	0,01761083	0,00407396
k__Bacteria;p__Spirochaetes	0,02821655	0,00653917	0,00784742	0,00644505	0,00723034	0,01127787	0,00564086
k__Bacteria;p__WS1	0,0017551	0,00037014	0,00069534	0,00033921	0,00088535	0,00112779	0
k__Bacteria;p__Cyanobacteria	0,00121507	0,00481184	0,00596007	0,00118725	0,00236093	0,00078078	0,00313381
k__Bacteria;p__OP3	0,01566086	0,00468846	0,00536406	0,01085482	0,00782057	0,02212197	0,00125353
k__Bacteria;p__OP8	0,01174565	0,00629241	0,00556273	0,0050882	0,00250848	0,00650646	0,00031338
k__Bacteria;p__Caldithrix	0,00621034	0,01443553	0,00645674	0,00474898	0,00575476	0,00858853	0,00376058
k__Bacteria;p__Chlorobi	0,00594033	0,00481184	0,00377471	0,00576662	0,00324627	0,00537868	0,0034472
k__Archaea;p__Euryarchaeota	0,00499527	0,02874769	0,00198669	0,01136364	0,00088535	0,01119112	0
k__Bacteria;p__Fusobacteria	0,00013501	0,00567551	0,00307937	0,00050882	0,00044267	8,6753E-05	0,00094014
k__Bacteria;p__Tenericutes	0	0,00074028	0,00447005	0,00067843	0,00191825	0	0,00031338
k__Bacteria;p__GN04	0,00243013	0,00296114	0,00238403	0,00339213	0,00044267	0,00529192	0,00532748
k__Bacteria;p__AC1	0,00499527	0,0032079	0,00089401	0,00220488	0,00088535	0,00694023	0
k__Bacteria;p__Elusimicrobia	0,00324018	0,00098705	0,00079468	0,00050882	0,00236093	0,00251583	0,00376058
k__Bacteria;p__TM6	0,00067504	0,00209747	0,00367537	0,00169607	0,00118046	0,00095428	0,00062676
k__Bacteria;p__Lentisphaerae	0,00108006	0,00024676	0,00099334	0,00067843	0,00132802	0,00156155	0,00720777
k__Bacteria;p__KSB3	0,00054003	0,00197409	0,00119201	0,00033921	0,00059023	8,6753E-05	0,00188029
k__Bacteria;p__NKB19	0,00081004	0,00037014	0,00109268	0,00135685	0,00014756	0,00320986	0,00031338
k__Bacteria;p__BRC1	0,00054003	0,00123381	0,00119201	0,00084803	0,00059023	0,0016483	0,00031338
k__Bacteria;p__	0,00094505	0,00024676	9,9334E-05	0,00050882	0,00014756	0,00078078	0,00062676
k__Bacteria;p__NC10	0,00013501	0	9,9334E-05	0,00491859	0	0,00156155	0
k__Bacteria;p__GN02	0,00081004	0,00123381	0,00079468	0,00101764	0	0,00078078	0,00219367
k__Bacteria;p__Fibrobacteres	0,00054003	0,00098705	9,9334E-05	0	0,00044267	0,00052052	0
k__Bacteria;p__SC4	0,00081004	0	0	0,00033921	0	0,00095428	0
k__Bacteria;p__[Caldithrix]	0,00040502	0,00222085	9,9334E-05	0,00016961	0	8,6753E-05	0
k__Bacteria;p__LCP-89	0,00081004	0,00111043	0,00059601	0,00033921	0,00044267	0,00034701	0,00031338
k__Bacteria;p__WS2	0,00081004	0,00024676	0,000298	0,00050882	0,00059023	0,00078078	0,00250705

Echantillon / Sample No.	37	70	54	72	59	57	64
Taxon de P1otu_table_L2	P2M32	P2M21	P2M6	P2M36	P2M38	P2M10	P2M16
k__Bacteria;p__Proteobacteria	0,5030813	0,71035699	0,53982535	0,69778819	0,55963303	0,5953132	0,51853643
k__Bacteria;p__Chloroflexi	0,05807063	0,02179328	0,04895475	0,02632123	0,01209341	0,04281208	0,06189555
k__Bacteria;p__Bacteroidetes	0,06577388	0,01068908	0,0796507	0,12650727	0,19349458	0,09553853	0,06511928
k__Bacteria;p__Planctomycetes	0,05854468	0,05759651	0,04498545	0,02039551	0,04587156	0,02208202	0,04158607
k__Bacteria;p__Acidobacteria	0,04313818	0,06174761	0,03069595	0,01708813	0,02835696	0,03695358	0,02756286
k__Bacteria;Other	0,01671012	0,0115193	0,02566817	0,00482326	0,01084237	0,00811176	0,03223727
k__Bacteria;p__Firmicutes	0,02156909	0,00041511	0,06747817	0,00716599	0,01251043	0,02433529	0,025951
k__Bacteria;p__Gemmatimonadetes	0,05427826	0,0143213	0,01905266	0,01329842	0,02627189	0,04055881	0,02836879
k__Bacteria;p__Actinobacteria	0,00888836	0,08146534	0,0195819	0,01047337	0,04753962	0,02884182	0,0070922
k__Bacteria;p__Verrucomicrobia	0,01659161	0,01079286	0,0351945	0,01391856	0,01125938	0,02298333	0,0248227
k__Bacteria;p__Nitrospirae	0,02915383	0,00269822	0,00661551	0,00344519	0,01000834	0,02478594	0,04045777
k__Bacteria;p__WS3	0,020621	0,00332088	0,01005557	0,00571901	0,00333611	0,00856242	0,00934881
k__Bacteria;p__Spirochaetes	0,01078455	0,00020756	0,00926171	0,00502997	0,00083403	0,00180261	0,01450677
k__Bacteria;p__WS1	0,0010666	0	0,00052924	6,8904E-05	0	0	0,00032237
k__Bacteria;p__Cyanobacteria	0,00319981	0,00643421	0,01058481	0,02191139	0,00125104	0,01036503	0,0046744
k__Bacteria;p__OP3	0,01173264	0,00051889	0,00502779	0,00144698	0,00083403	0,00630915	0,01434558
k__Bacteria;p__OP8	0,00485897	0	0,00317544	0,00110246	0	0,00090131	0,00483559
k__Bacteria;p__Caldithrix	0,01528798	0,00083022	0,00185234	0,0037208	0,00083403	0,00495719	0,02836879
k__Bacteria;p__Chlorobi	0,00734771	0,00020756	0,00476316	0,00282505	0,00125104	0,00270392	0,01192779
k__Archaea;p__Euryarchaeota	0,00545153	0,00031133	0	0,00117136	0,00250209	0,00045065	0,00257898
k__Bacteria;p__Fusobacteria	0,00865134	0,00010378	0,01481874	0,00661476	0,01042535	0,00090131	0,0046744
k__Bacteria;p__Tenericutes	0	0,00020756	0,00449854	0,00289396	0,01668057	0,00540784	0,00128949
k__Bacteria;p__GN04	0,00284428	0,00020756	0,0013231	0,00213602	0,00041701	0,00135196	0,0082205
k__Bacteria;p__AC1	0,00213321	0	0	0,00013781	0	0	0,00145068
k__Bacteria;p__Elusimicrobia	0,00189618	0,00041511	0,00079386	0,00034452	0	0,00225327	0,00161186
k__Bacteria;p__TM6	0,00225172	0,00020756	0,00158772	0,00013781	0	0	0,00145068
k__Bacteria;p__Lentisphaerae	0,00082958	0,00020756	0,00714475	0,00027561	0	0,00090131	0,0011283
k__Bacteria;p__KSB3	0,00094809	0,00010378	0,00079386	0,00027561	0	0,00090131	0,00128949
k__Bacteria;p__NKB19	0,00296279	0,00010378	0,00026462	0,00041342	0,00041701	0,00135196	0,00193424
k__Bacteria;p__BRC1	0,00213321	0,00010378	0,00211696	0,00041342	0,00041701	0,00180261	0,0011283
k__Bacteria;p__	0,00047405	0	0	6,8904E-05	0	0,00045065	0,00016119
k__Bacteria;p__NC10	0,00888836	0	0	0	0	0	0
k__Bacteria;p__GN02	0,00165916	0,00031133	0,00052924	0,00020671	0,00041701	0	0,00048356
k__Bacteria;p__Fibrobacteres	0,00035553	0	0	0,00013781	0	0,00270392	0,00177305
k__Bacteria;p__SC4	0,0010666	0	0	0,00013781	0	0,00090131	0,00096712
k__Bacteria;p__[Caldithrix]	0,00130363	0	0,00079386	0	0	0	0,00016119
k__Bacteria;p__LCP-89	0,00035553	0	0,00026462	0,00013781	0	0	0,00032237
k__Bacteria;p__WS2	0,00059256	0	0	0,00013781	0	0,00045065	0,00177305

Echantillon / Sample No.	39	52	
Taxon de P1otu_table_L2	P2M33	P2M4	Mean
k__Bacteria;p__Proteobacteria	0,5311973	0,7141669	0,57382427
k__Bacteria;p__Chloroflexi	0,1163575	0,02827835	0,08651571
k__Bacteria;p__Bacteroidetes	0,05986509	0,0587746	0,047502
k__Bacteria;p__Planctomycetes	0,04300169	0,02107014	0,03887901
k__Bacteria;p__Acidobacteria	0,01602024	0,02328805	0,03166854
k__Bacteria;Other	0,00590219	0,00803992	0,0279943
k__Bacteria;p__Firmicutes	0,05986509	0,01829775	0,02749839
k__Bacteria;p__Gemmatimonadetes	0,00590219	0,01413917	0,02342886
k__Bacteria;p__Actinobacteria	0,01686341	0,01524813	0,02138269
k__Bacteria;p__Verrucomicrobia	0,02192243	0,01164403	0,01707797
k__Bacteria;p__Nitrospirae	0,0143339	0,00277239	0,01547579
k__Bacteria;p__WS3	0,01686341	0,00914888	0,00996223
k__Bacteria;p__Spirochaetes	0,01011804	0,00443582	0,00853959
k__Bacteria;p__WS1	0,00084317	0	0,00751972
k__Bacteria;p__Cyanobacteria	0,02613828	0,00332686	0,00739131
k__Bacteria;p__OP3	0,00168634	0,00221791	0,00692884
k__Bacteria;p__OP8	0,00505902	0,00221791	0,00668996
k__Bacteria;p__Caldithrix	0,01686341	0,00304963	0,00595964
k__Bacteria;p__Chlorobi	0,00505902	0,00194067	0,00454687
k__Archaea;p__Euryarchaeota	0	0,00027724	0,00350032
k__Bacteria;p__Fusobacteria	0,00421585	0,00138619	0,00301284
k__Bacteria;p__Tenericutes	0,00084317	0,04685334	0,00281325
k__Bacteria;p__GN04	0,00252951	0,00166343	0,00215136
k__Bacteria;p__AC1	0,00337268	0,00027724	0,00189013
k__Bacteria;p__Elusimicrobia	0	0,00055448	0,00154158
k__Bacteria;p__TM6	0,00168634	0,00083172	0,00132668
k__Bacteria;p__Lentisphaerae	0	0,00138619	0,00124915
k__Bacteria;p__KSB3	0,00252951	0	0,00123142
k__Bacteria;p__NKB19	0	0,00055448	0,00105162
k__Bacteria;p__BRC1	0,00084317	0,00027724	0,00101233
k__Bacteria;p__	0	0	0,00072308
k__Bacteria;p__NC10	0	0	0,00068466
k__Bacteria;p__GN02	0	0,00110895	0,00063093
k__Bacteria;p__Fibrobacteres	0	0,00027724	0,00061747
k__Bacteria;p__SC4	0	0	0,00055878
k__Bacteria;p__[Caldithrix]	0	0	0,00055582
k__Bacteria;p__LCP-89	0	0,00055448	0,00054064
k__Bacteria;p__WS2	0,00084317	0,00083172	0,00052236

Physical-Chemical Chilika Lake (Chapter 4)

ample Numbe	Sample Name	ate of Sampli	Sampling Season	/ % DEP week	on (as per CDA	pH
No.1	CP-S	15-avr-11	SUMMER	35.5 / 15	19. 71 556 N	9,32
No.2	CP-R	20-juil-11	RAINY	218.1 / 30		8,1
No.3	CP-W	18-janv-12	WINTER	25.1 / 149	85.19697 E	9,47
No. 2 *	Post Phailin	02-nov-13	ilin (after 12th Oct, 2013)			7,2
No.4	CH-S	15-avr-11	SUMMER	35.5 / 15	19.71292 N	9,38
No.5	CH-R	20-juil-11	RAINY	218.1 / 30		8,38
No.6	CH-W	18-janv-12	WINTER	25.1 / 149	85. 20828 E	8,93
No. 5 *	Post Phailin	02-nov-13	ilin (after 12th Oct, 2013)			8,48
No.61	NI2-S	15-avr-11	SUMMER	35.5 / 15	19. 69371 N	9,35
No.62	NI2-R	12-sept-11	RAINY	204.2 / 19		8,42
No.63	NI2-W	19-janv-12	WINTER	25.1 / 149	85. 29115 E	9,21
No. 62 *	Post Phailin	30-oct-13	ilin (after 12th Oct, 2013)			7,78
No.7	NO2-S	15-avr-11	SUMMER	35.5 / 15	19. 71208 N	9,4
No.8	NO2-R	12-sept-11	RAINY	204.2 / 19		8,38
No.9	NO2-W	19-janv-12	WINTER	25.1 / 149	85. 32519 E	8,9
No. 8 *	Post Phailin	30-oct-13	ilin (after 12th Oct, 2013)			7,91
No.64	NI5-S	15-avr-11	SUMMER	35.5 / 15	19.68157 N	9,25
No.65	NI5-R	12-sept-11	RAINY	204.2 / 19		8,22
No.66	NI5-W	19-janv-12	WINTER	25.1 / 149	85.30440 E	9,12
No. 65 *	Post Phailin	30-oct-13	ilin (after 12th Oct, 2013)			7,87
No.10	NO5-S	15-avr-11	SUMMER	35.5 / 15	19.67914 N	8,55
No.11	NO5-R	12-sept-11	RAINY	204.2 / 19		8,34
No.12	NO5-W	19-janv-12	WINTER	25.1 / 149	85.31046 E	9,03
No. 11 *	Post Phailin	30-oct-13	ilin (after 12th Oct, 2013)			8,02
No.13	KJ-S	15-avr-11	SUMMER	35.5 / 15	19. 66649 N	8,6
No.14	KJ-R	20-juil-11	RAINY	218.1 / 30		8,21
No.15	KJ-W	18-janv-12	WINTER	25.1 / 149	85. 21780 E	8,72
No. 14 *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)			8,55
No.22	HB-S	25-mai-11	SUMMER	7.6 / 89	19. 82965 N	7,8
No.23	HB-R	03-sept-11	RAINY	204.2 / 19		8,18
No.24	HB-W	15-févr-12	WINTER	0.0 / 100	85. 40778 E	8,55
No. 23 *	Post Phailin	19-nov-13	ilin (after 12th Oct, 2013)			7,58
No.19 a	BM I (0 mtr)-S	25-mai-11	SUMMER	7.6 / 89	19. 78224 N	9,1
No.19 b	BM II (50 mtrs)-S	25-mai-11	SUMMER	7.6 / 89		9
No.19 c	M III (100 mtrs)-	25-mai-11	SUMMER	7.6 / 89	85. 30041 E	9,15
No.20 a	BM I (0 mtr)-R	03-sept-11	RAINY	204.2 / 19	19. 78224 N	8,4
No.20 b	BM II (50 mtrs)-R	03-sept-11	RAINY	204.2 / 19		8,6
No.20 c	M III (100 mtrs)-	03-sept-11	RAINY	204.2 / 19	85. 30041 E	8,4
No.21 a	BM I (0 mtr)-W	15-févr-12	WINTER	0.0 / 100	19. 78224 N	9,55
No.21 b	BM II (50 mtrs)-W	15-févr-12	WINTER	0.0 / 100		9,72
No.21 c	M III (100 mtrs)-V	15-févr-12	WINTER	0.0 / 100	85. 30041 E	10,1

Sample	Sample Name	Date	Sampling Season	) / % DEP week I	Loc	pH
No.21 c	BM III (100 mtrs)-W	15-févr-12	WINTER	0.0 / 100	85. 30041 E	10,1
No.20 a *	Post Phailin	19-nov-13	ilin (after 12th Oct, 2013)		19. 78224 N	
No.20 b *	Post Phailin	19-nov-13	ilin (after 12th Oct, 2013)			
No.20 c *	Post Phailin	19-nov-13	ilin (after 12th Oct, 2013)		85. 30041 E	9,35
No.16	KG-S	16-avr-11	SUMMER	35.5 / 15	19. 84699 N	6,85
No.17	KG-R	20-août-11	RAINY	297.2 / 12		7,7
No.18	KG-W	14-févr-12	WINTER	0.0 / 100	85. 40778 E	8,63
No.17 *	Post Phailin	19-nov-13	ilin (after 12th Oct, 2013)			7,46
No.25	CDA18-S	16-avr-11	SUMMER	35.5 / 15		8,4
No.26	CDA18-R	20-août-11	RAINY	297.2 / 12	19.826590 N	8,84
No.27	CDA18-W	14-févr-12	WINTER	0.0 / 100		8,91
No. 26 *	Post Phailin	02-nov-13	ilin (after 12th Oct, 2013)		85.474830 E	8,1
No.28	CDA19-S	16-avr-11	SUMMER	35.5 / 15	19.783380 N	9,12
No.29	CDA19-R	20-août-11	RAINY	297.2 / 12		8,62
No.30	CDA19-W	14-févr-12	WINTER	0.0 / 100	85.492870 E	9,02
No. 29 *	Post Phailin	02-nov-13	ilin (after 12th Oct, 2013)			7,87
No.31	INS-S	25-mai-11	SUMMER	36.9 / 57	19. 67853 N	8,49
No.32	INS-R	03-sept-11	RAINY	160.2 / 24		7,8
No.33	INS-W	14-févr-12	WINTER	1.0 / 96	85. 19349 E	9,43
No. 32 *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)			8,15
No.37 a	GS I (0 mtr)-S	25-mai-11	SUMMER	36.9 / 57	19. 51626 N	8,69
No.37 b	GSII (50 mtrs)-S	25-mai-11	SUMMER	36.9 / 57		8,58
No.37 c	GSIII (100 mtrs)-S	25-mai-11	SUMMER	36.9 / 57	85. 12477 E	8,6
No.38 a	GSI (0 mtr)-R	03-sept-11	RAINY	160.2 / 24	19. 51626 N	8,25
No.38 b	GSII (50 mtrs)-R	03-sept-11	RAINY	160.2 / 24		8,27
No.38 c	GSIII (100 mtrs)-R	03-sept-11	RAINY	160.2 / 24	85. 12477 E	8,19
No.39 a	GSI (0 mtr)-W	14-févr-12	WINTER	1.0 / 96	19. 51626 N	8,87
No.39 b	GSII (50 mtrs)-W	14-févr-12	WINTER	1.0 / 96		8,95
No.39 c	GS III (100 mtrs)-W	14-févr-12	WINTER	1.0 / 96	85. 12477 E	8,83
No.38 a *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)		19. 51626 N	
No.38 b *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)			
No.38 c *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)		85. 12477 E	8,06
No.34	PK-S	15-avr-11	SUMMER	68.3 / 65		8,62
No.35	PK-R	20-juil-11	RAINY	184.8 / 16	19.632150 N	8,22
No.36	PK-W	18-janv-12	WINTER	53.2 / 300		8,73
No. 35 *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)		85.210660 E	7,69
No.40	PL-S	15-avr-11	SUMMER	68.3 / 65	19. 55549 N	8,64
No.41	PL-R	20-juil-11	RAINY	184.8 / 16		8,17
No.42	PL-W	18-janv-12	WINTER	53.2 / 300	85. 18265 E	8,91
No. 41 *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)			8,49
No.67	HI-S	15-avr-11	SUMMER	68.3 / 65	19. 58985 N	8,52
No.68	HI-R	20-juil-11	RAINY	184.8 / 16		8,47
No.69	HI-W	18-janv-12	WINTER	53.2 / 300	85. 14725 E	8,81
No. 68 *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)			8,39

Sample	Name	Date	Season	/ % DEP week	Loc	pH
No.69	HI-W	18-janv-12	WINTER	53.2 / 300	85. 14725 E	8,81
No. 68 *	Post Phailin	25-nov-13	ilin (after 12th Oct, 2013)			8,39
No.43	MM-S	11-juin-11	SUMMER	156.9 / 7	19. 69540 N	8
No.44	MM-R	19-août-11	RAINY	259.8 / 21		8,66
No.45	MM-W	19-janv-12	WINTER	6.7 / 58	85. 42625 E	8,92
No. 44 *	Post Phailin	01-nov-13	ilin (after 12th Oct, 2013)			8,27
No.46	GM-S	11-juin-11	SUMMER	156.9 / 7	19. 69042 N	8,3
No.47	GM-R	19-août-11	RAINY	259.8 / 21		8,48
No.48	GM-W	19-janv-12	WINTER	6.7 / 58	85. 41981 E	9,2
No. 47 *	Post Phailin	02-nov-13	ilin (after 12th Oct, 2013)			7,87
No.49	CN-S	11-juin-11	SUMMER	156.9 / 7	19. 65565 N	8,3
No.50	CN-R	19-août-11	RAINY	259.8 / 21		8,41
No.51	CN-W	19-janv-12	WINTER	6.7 / 58	85. 45972 E	8,72
No. 50 *	Post Phailin	01-nov-13	ilin (after 12th Oct, 2013)			8,12
No.52	OM-S	11-juin-11	SUMMER	156.9 / 7	19.731170 N	8,2
No.53	OM-R	19-août-11	RAINY	259.8 / 21		8,34
No.54	OM-W	19-janv-12	WINTER	6.7 / 58		9,5
No. 53 *	Post Phailin	01-nov-13	ilin (after 12th Oct, 2013)		85.650640 E	8,42
No.70	AK-S	11-juin-11	SUMMER	156.9 / 7	19. 70535 N	8,2
No.71	AK-R	19-août-11	RAINY	259.8 / 21		8,24
No.72	AK-W	19-janv-12	WINTER	6.7 / 58	85. 58926 E	9,15
No. 71 *	Post Phailin	01-nov-13	ilin (after 12th Oct, 2013)			8,36
No.55	BK-S	11-juin-11	SUMMER	156.9 / 7	19. 68452 N	8,2
No.56	BK-R	19-août-11	RAINY	259.8 / 21		8,19
No.57	BK-W	19-janv-12	WINTER	6.7 / 58	85. 52730 E	8,65
No. 56 *	Post Phailin	01-nov-13	ilin (after 12th Oct, 2013)			7,23
No.58	RB-S	11-juin-11	SUMMER	156.9 / 7	19. 66121 N	8,2
No.59	RB-R	19-août-11	RAINY	259.8 / 21		8,28
No.60	RB-W	19-janv-12	WINTER	6.7 / 58	85. 49887 E	8,7
No. 59 *	Post Phailin	01-nov-13	ilin (after 12th Oct, 2013)			7,17

Sample	TURBIDITY	SALINITY	CONDUCTIVITY	COLOUR	AIR TEMP.	WATER TEMP.
No.1	11,2	7,5	12,1	Transparent Green	30,5	30,5
No.2	18	7,9	12,6	Transparent Green	28,9	29,8
No.3	31,5	7,6	12,6	Transparent Green	24,2	25,5
No. 2 *	12,6	2,8	4,94	Transparent Green	27	29,3
No.4	12,5	8,6	13,8	Transparent Green	30	30,2
No.5	17,3	8	13	Turbid green	27,9	29,1
No.6	27	8,2	13,5	Transparent Green	22,2	22,4
No. 5 *	12,5	3,1	5,47	Transparent Green	25,2	27,6
No.61	14,5	18,2	27,2	Transparent Green	29,7	30,7
No.62	23,8	4	6,8	Turbid green	30,3	30,8
No.63	29,5	7,4	12,1	Transparent Green	25	24
No. 62 *	50	2,6	4,6	Transparent Green	25	27,7
No.7	18	18,3	27,3	Transparent Green	28,3	30,7
No.8	51,3	4,2	8,02	Transparent Green	30	29,8
No.9	46	7,9	13	Transparent Green	25,6	23,2
No. 8 *	53	3,3	5,35	Transparent Green	25	27,2
No.64	16,5	18,4	27,3	Transparent Green	30,4	32,5
No.65	17,6	8,6	13,8	Turbid green	31,3	32
No.66	40,8	8,2	13,4	Transparent Green	24,6	26,4
No. 65 *	80	3,8	6,48	Turbid green	25,5	28,2
No.10	15,2	18,2	27,1	Turbid green	28,2	31,3
No.11	15,8	9,3	14,6	Transparent Green	31	31,3
No.12	40,2	7,1	11,7	Turbid green	27	25,6
No. 11 *	22,7	3,3	3,77	Turbid green	25,4	27,8
No.13	85	14,2	21,6	Turbid green	28,4	30,1
No.14	26,2	15,9	24,1	Turbid green	28,3	29,8
No.15	41	9,1	14,8	Turbid green	21,2	22,2
No. 14 *	17,5	3,8	6,63	Transparent Green	23,6	24,6
No.22	24,3	8,8	14	Turbid green	33,3	33,7
No.23	68,1	0	30,1	Turbid green	29	31
No.24	85	4,4	7,38	Turbid green	26,9	27,1
No. 23 *	61,9	2,8	0,503	Turbid green	22,9	24,6
No.19 a	18,7	8,8	14	Turbid green	31,8	32,6
No.19 b	18,6	8,8	14	Turbid green	31,6	32,6
No.19 c	30,9	8,9	14	Turbid green	32	32,4
No.20 a	16,7	3,3	5,78	Turbid green	30,4	29,3
No.20 b	21,3	1	1,96	Turbid green	30,4	29,8
No.20 c	16,7	3,3	5,78	Turbid green	30,2	29,8
No.21 a	51,3	6,6	10,8	Turbid green	26,7	27,9
No.21 b	20,6	6,6	10,8	Turbid green	27,1	28,4
No.21 c	18,9	6,6	10,8	Turbid green	28,2	28,6
No.20 a *						
No.20 b *						
No.20 c *	38,7	3,3	0,575	Transparent Green	23	24,6

Samples	TURBIDITY	SALINITY	CONDUCTIVITY	COLOUR	AIR TEMP.	WATER TEMP.
No.16	48,4	8,9	14	Turbid green	31,6	32,8
No.17	90	0,1	24,8	Turbid green	26,4	28,9
No.18	171	1,6	2,97	Turbid green	24,7	27,7
No.17 *	127,1	1,1	0,207	Turbid green	23	24,2
No.25	90	16,4	24,6	Transparent Gree	31,3	34
No.26	41,5	0,1	2,44	Transparent Gree	28,2	32
No.27	102,5	1,1	2,12	Transparent Gree	25,3	27,7
No. 26 *	34	0	0,09	Turbid green	23,6	27,8
No.28	109	1,5	2,67	Turbid green	29	30,6
No.29	163	0,1	2,72	Turbid green	28,5	30,1
No.30	19,2	9,7	15,6	Turbid green	24,2	25,5
No. 29 *	186	0,1	0,194	Turbid green	23,9	26,8
No.31	18,6	16,9	25,3	Transparent Gree	29	31,5
No.32	32,5	7,8	12,4	Transparent Gree	28,6	29,1
No.33	65	9,6	15,3	Transparent Gree	26,1	29,4
No. 32 *	20,4	3,9	6,78	Transparent Gree	24,5	25,6
No.37 a	19,7	16,2	24,3	Turbid green	31,5	34,3
No.37 b	29,4	16,1	24,3	Turbid green	31,4	32,1
No.37 c	40	16,1	24,2	Turbid green	30,6	32,2
No.38 a	15,8	12,4	19,3	Turbid green	26,7	29,1
No.38 b	18,4	12	18,7	Turbid green	26,7	28,6
No.38 c	6,7	11,9	18,2	Turbid green	26,7	29
No.39 a	44	13,4	20,9	Turbid green	26,9	27
No.39 b	20,1	13,2	20,5	Turbid green	27,1	27,4
No.39 c	27,9	13,2	20,6	Turbid green	26,3	26,7
No.38 a *						
No.38 b *						
No.38 c *	17	4,9	8,32	Turbid green	24,5	26,2
No.34	27,8	13	20,1	Turbid green	29,5	31,3
No.35	18	13,5	20,8	Turbid green	27,9	29,6
No.36	32	9,7	15,7	Turbid green	21,7	22,1
No. 35 *	16	4,1	7,1	Transparent Gree	23,6	24,3
No.40	21,5	13,6	20,8	Turbid green	29,6	31,2
No.41	11,8	13	20,1	Turbid green	28,7	29,7
No.42	35	11,5	18,3	Turbid green	22,5	22,9
No. 41 *	18	4,6	7,6	Transparent Gree	24,6	25
No.67	34	13,6	20,9	Transparent Gree	30,5	31,9
No.68	24,2	15,1	22,9	Transparent Gree	30,7	31
No.69	46	10,6	17	Transparent Gree	23	23,4
No. 68 *	18	4,7	7,98	Transparent Gree	24,5	26
No.43	34,1	29,9	44,9	Turbid green	29,2	32
No.44	18,5	5,4	18,9	Turbid green	26	29
No.45	80	14,8	23,1	Turbid green	23,1	25
No. 44 *	74	1,2	2,25	Turbid green	24	26,8

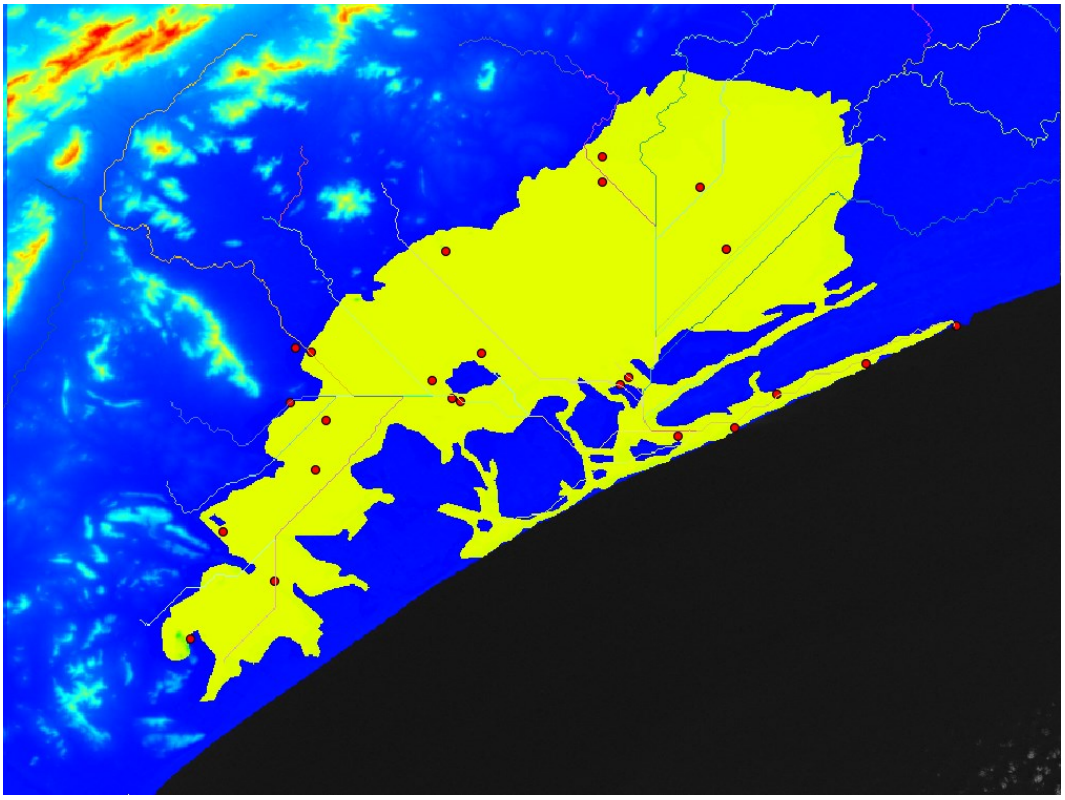
Sample	TURBIDITY	SALINITY	CONDUCTIVITY	COLOUR	AIR TEMP.	WATER TEMP.
No.46	42,8	30,6	44,7	Turbid green	29,5	33
No.47	46	6,7	10,9	Turbid green	27,1	28,8
No.48	46	18,4	28,1	Turbid green	23,5	25,5
No. 47 *	163	0,6	1,31	Turbid green	23	26,5
No.49	32	32,3	45,9	Turbid green	29,6	28,7
No.50	90	7,2	11,7	Turbid green	26	29,8
No.51	37	20,9	31,4	Turbid green	23,8	23,8
No. 50 *	109	1,2	2,34	Turbid green	24,5	26,9
No.52	16,7	28,2	27,5	Turbid green	28,7	31,4
No.53	19	17,6	26,4	Turbid green	26,7	29,9
No.54	33,9	20,4	30,7	Turbid green	23,7	25,7
No. 53 *	66	0,4	0,95	Turbid green	25,5	28
No.70	40,3	31,3	38,5	Turbid green	29	30,6
No.71	39,2	22,4	32,8	Turbid green	26,8	29,7
No.72	43,7	22,6	33,8	Turbid green	23,4	24,1
No. 71 *	80	1	1,87	Turbid green	25,1	28,2
No.55	36,8	32,2	45,9	Turbid green	28,3	30
No.56	22,7	27,5	40,4	Turbid green	26,3	28,7
No.57	34	24,5	36,4	Turbid green	23	24,1
No. 56 *	80	1,7	2,9	Turbid green	25,1	28,4
No.58	32,4	32,3	46,1	Turbid green	28,4	28,5
No.59	33,3	30,3	43,3	Turbid green	26	29,1
No.60	132	23,3	34,6	Turbid green	24,2	24,3
No. 59 *	65	2	3,77	Turbid green	24,9	27,8

Sample	DO	Nitrite	Nitrate	Ammonia	Phosphate	Total Iron	TOC
No.1	5,4	3,3	10,64	206,409	1,35	0,735	3
No.2	4,03	1,19	1,88	62,823	1,354	0,374	5,4
No.3	4,70	0,677	1,06	46,873	0,547	0,444	4,3
No. 2 *	6,78	0,18	3,76	152,9	0,32	0,52	5,3
No.4	6,6	3,9	12,07	117,954	2,58	0,557	6
No.5	5,31	1,29	2,03	9,994	2,58	0,273	8,2
No.6	9,85	0,965	4,52	56,723	1,690	0,758	6,2
No. 5 *	10,7	0,1	3,33	126,85	1,45	0,529	6,5
No.61	7,5	8,6	10,45	51,15	3,10	1,131	3,48
No.62	7,64	0,27	1,35	90,659	0,136	0,217	5,2
No.63	6,62	1,355	4,87	16,976	1,733	0,729	7,9
No. 62 *	8,1	0,45	0,52	19,19	0,47	0,7	5,43
No.7	7,35	7,52	9,42	27,925	2,87	1,147	11,81
No.8	7,8	0,35	2,35	24,084	0,114	0,279	7,5
No.9	6,74	1,28	5,64	62,077	1,608	0,927	7
No. 8 *	8,8	0,53	0,58	97,568	0,67	0,81	8,2
No.64	7,92	8,1	9,52	55,213	2,98	0,68	4,72
No.65	7,74	0,31	1,87	44,359	0,218	0,918	3,2
No.66	6,65	1,31	5,31	117,654	1,712	0,437	3,9
No. 65 *	8,65	0,43	0,54	178,304	1,01	0,98	5,8
No.10	7,43	7,41	9,8	41,203	2,54	0,364	6,75
No.11	6,9	0,51	1,75	39,729	0,173	0,449	5,6
No.12	6,67	1,37	5,25	10,926	1,681	0,266	8,4
No. 11 *	9,16	0,49	0,61	124,63	0,724	0,36	6,75
No.13	6,6	4,6	12,08	40,692	2,32	0,608	6,34
No.14	6,81	1,64	2,87	35,628	2,322	0,315	2,57
No.15	7,29	1,108	5,05	24,322	1,505	0,484	6,2
No. 14 *	11,2	0,08	0,77	156,24	1,06	0,47	5,24
No.22	7,2	1,40	2,31	17,212	1,39	1,301	8,05
No.23	7,33	0,31	0,82	17,197	0,047	1,3	7,02
No.24	7,26	0,513	1,18	16,49	0,73	0,235	5,3
No. 23 *	5,7	1,57	1,63	228,19	0,89	0,95	6,83
No.19 a	7,0	1,89	8,53	147,5	1,11	1,009	2,57
No.19 b	6,74	0,41	7,29	125,8	1,218	0,701	1,68
No.19 c	6,8	0,48	7,87	97,23	1,31	0,924	2,12
No.20 a	7,39	0,35	2,01	69,962	0,136	0,214	7,1
No.20 b	7,27	0,38	1,81	56,25	0,172	0,15	4,091
No.20 c	7,4	0,4	1,51	38,14	0,153	0,25	6,51
No.21 a	6,66	0,431	1,59	117,698	0,64	0,356	8,9
No.21 b	6,94	0,532	2,27	25,275	0,732	1,469	5,6
No.21 c	7,24	0,52	1,94	24,86	0,69	0,254	5,5
No.20 a *							
No.20 b *							

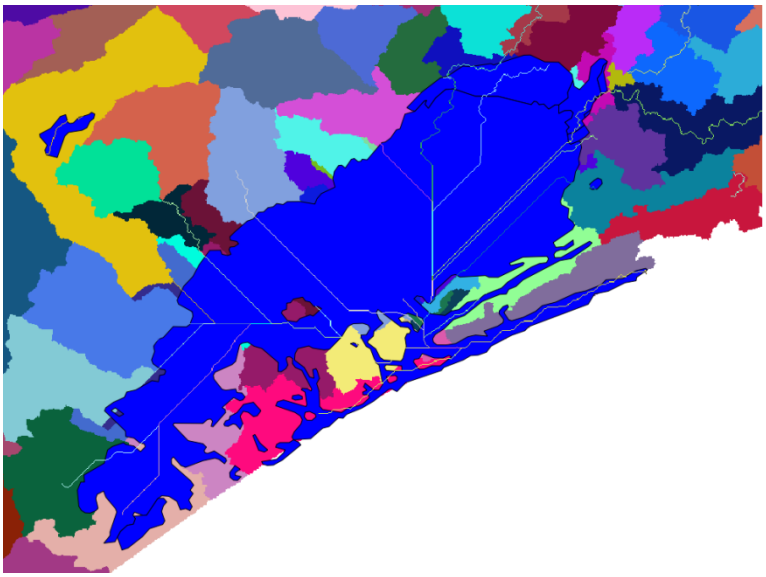
Sample	DO	Nitrite	Nitrate	Ammonia	Phosphate	Total Iron	TOC
No.20 c *	5,8	0,26	4,7	196,85	1,25	0,67	5,41
No.16	8,3	6,3	2,76	22,488	0,65	0,212	1,5
No.17	5,44	2,83	4,25	49,973	1,7	1,181	6,6
No.18	6,05	1,006	3,63	95,615	0,64	0,902	5,6
No.17 *	7,2	0,93	0,77	50,16	0,45	0,75	4,7
No.25	7,7	0,8	15,66	54,81	2,45	0,42	5,62
No.26	10,82	1,99	3,93	71,081	1,61	0,201	2,3
No.27	9,12	0,903	1,67	40,982	0,87	0,604	5,1
No. 26 *	9,4	0,72	0,9	15,94	0,5	0,48	4,42
No.28	7,7	6,2	8,92	48,34	1,45	0,75	5,31
No.29	8,61	3,90	7,43	64,246	1,167	0,996	4,77
No.30	6,46	0,554	1,69	29,252	1,04	0,337	3,7
No. 29 *	11,2	1,36	1,27	19,11	0,46	0,7	4,84
No.31	7,3	2,33	15,85	37,569	1,70	1,053	0,45
No.32	6,53	0,27	0,94	11,85	0,091	0,313	5,5
No.33	7,81	0,369	1,28	149,191	0,64	0,814	8,9
No. 32 *	10,2	0,11	2,52	134,01	0,24	0,74	5,24
No.37 a	7,8	2,56	21,33	12,778	1,06	0,716	5,8
No.37 b	6,53	3,62	18,33	19,28	1,251	0,624	5,6
No.37 c	6,31	3,05	17,13	24,27	1,114	0,705	4,97
No.38 a	6,02	0,25	2,54	96,568	0,045	1,215	2,3
No.38 b	6,37	0,7	2,71	58,27	0,067	1,107	2,1
No.38 c	6,72	0,72	2,74	87,56	0,081	0,915	1,84
No.39 a	7,07	0,595	0,94	21,236	0,82	0,214	5,8
No.39 b	7,16	0,56	1,38	11,884	0,847	0,557	5,7
No.39 c	7,2	0,54	1,94	22,973	0,801	0,335	5,52
No.38 a *							
No.38 b *							
No.38 c *	8,7	0,08	0,83	144,03	0,14	0,64	4,42
No.34	6,7	8,2	25,67	49,915	2,13	0,968	0,75
No.35	5,50	1,33	1,59	84,24	2,129	0,45	2,6
No.36	8,61	1,17	3,56	73,118	1,916	0,664	3
No. 35 *	10,5	0,04	0,26	127,15	0,24	0,7	2,34
No.40	6,1	8,0	21,43	56,45	0,84	0,845	0,95
No.41	6,34	1,68	1,85	59,792	0,838	0,957	1,45
No.42	8,26	1,088	3,75	31,353	1,870	0,857	2,7
No. 41 *	12,0	0,08	9,65	118,73	0,35	0,9	1,83
No.67	7,2	5,0	10,31	65,32	2,32	0,74	1,31
No.68	6,88	1,79	3,17	67,112	2,32	1,388	1,36
No.69	8,06	1,129	5,75	26,986	2,140	0,211	5,7
No. 68 *	13,4	0,04	4,61	118,03	0,22	0,75	2,95
No.43	6,40	0,80	12,36	101,639	0,37	1,208	3,63
No.44	6,56	1,77	4,27	32,708	2,73	1,47	7,9

Sample	DO	Nitrite	Nitrate	Ammonia	Phosphate	Total Iron	TOC
No.45	7,94	1,416	5,73	24,799	2,281	0,249	5
No. 44 *	10,6	0,64	0,94	27,5	0,65	1,02	4,65
No.46	8,93	0,97	10,64	94,936	0,27	1,144	0,68
No.47	7,46	3,28	6,14	78,736	1,61	0,968	6,2
No.48	4,58	1,499	5,72	84,089	2,464	0,588	7,6
No. 47 *	11,0	1,02	1,86	39,17	1,42	1,62	4,93
No.49	8,42	0,53	5,56	21,165	0,27	1,223	2,72
No.50	5,82	2,75	5,92	18,392	2,23	0,485	7,6
No.51	7,49	1,601	5,88	50,056	1,916	0,172	7,8
No. 50 *	8,4	1,21	1,6	36,85	0,96	1,42	6,05
No.52	9,12	1,11	6,45	23,454	0,18	1,115	4,68
No.53	6,18	1,79	3,94	25,575	1,95	1,275	9,1
No.54	9,18	1,827	5,90	35,628	2,650	0,093	6,9
No. 53 *	9,4	0,05	1,44	29,18	0,33	1,25	6
No.70	7,84	0,72	7,40	7,319	0,41	1,218	0,3
No.71	5,76	3,88	2,93	31,078	1,8	0,22	8,4
No.72	7,62	1,786	5,75	55,101	2,650	0,272	9,3
No. 71 *	8,7	0,02	1,59	37,98	0,38	0,63	8,54
No.55	6,82	0,86	4,71	27,083	0,42	0,164	3,93
No.56	6,88	2,81	5,34	27,072	1,5	0,218	7,5
No.57	7,58	1,704	5,35	55,198	1,597	0,353	8,2
No. 56 *	8,4	0,08	1,95	41,17	0,36	0,34	4,6
No.58	8,06	1,05	5,73	132,373	0,27	1,195	9,99
No.59	6,59	2,34	5,40	36,764	2,1	0,256	8,7
No.60	6,50	1,334	5,61	63,593	1,551	0,372	5,5
No. 59 *	10,6	0,52	1,76	33,28	0,52	0,62	7,62

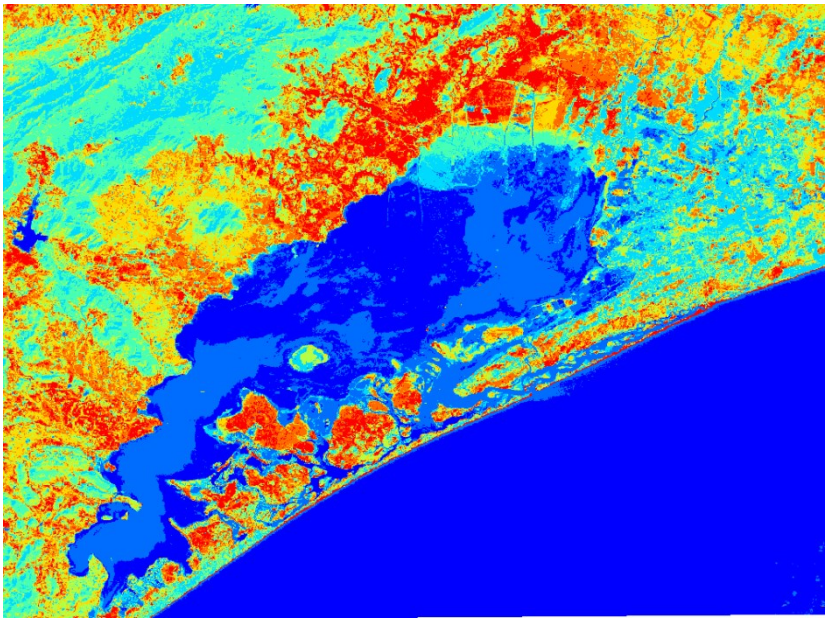
DEM



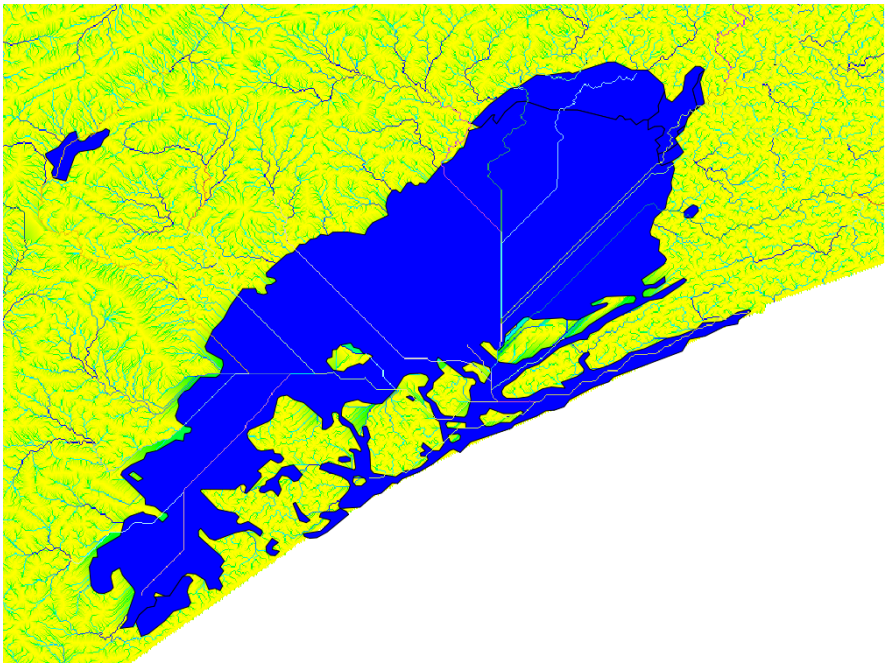
Watershed



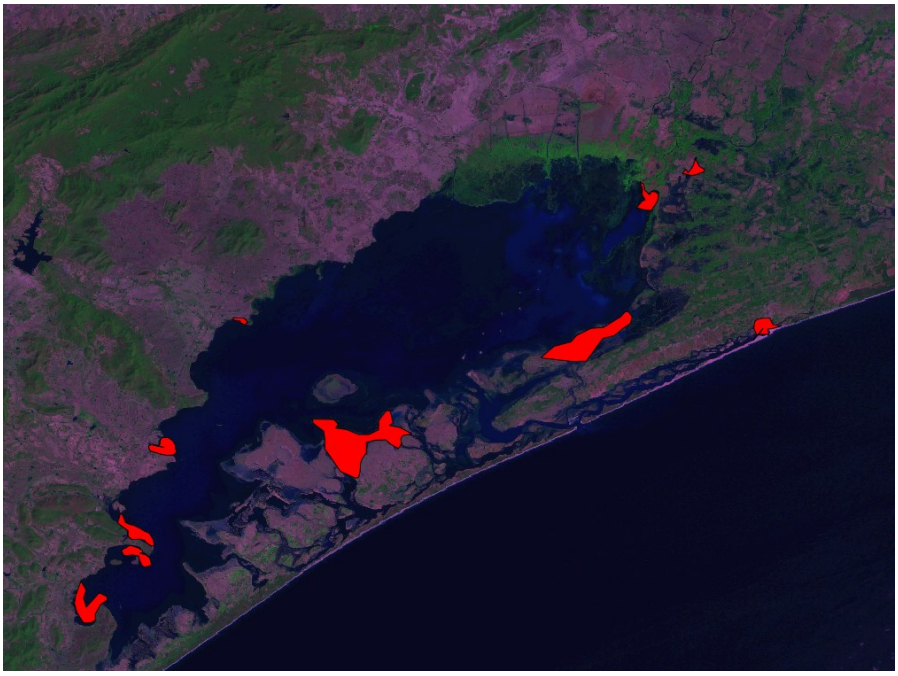
Classification



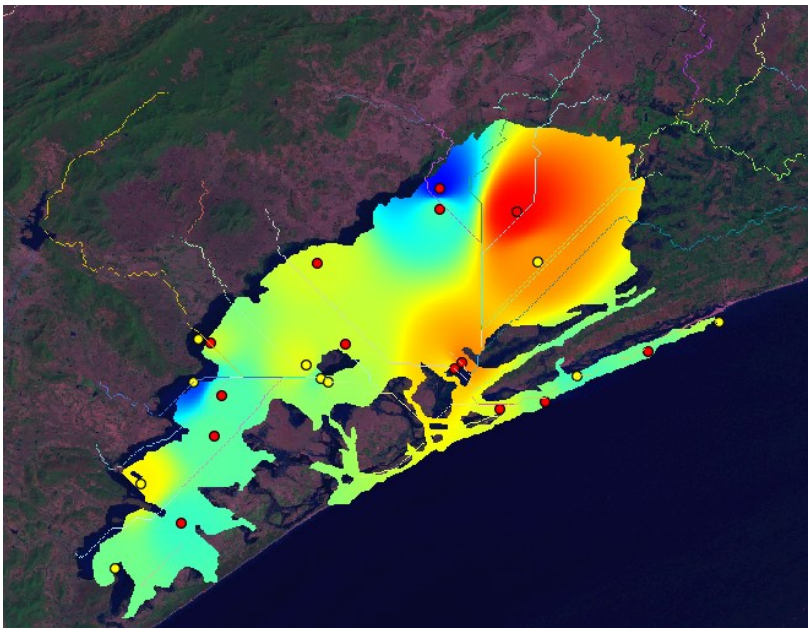
Accumulation



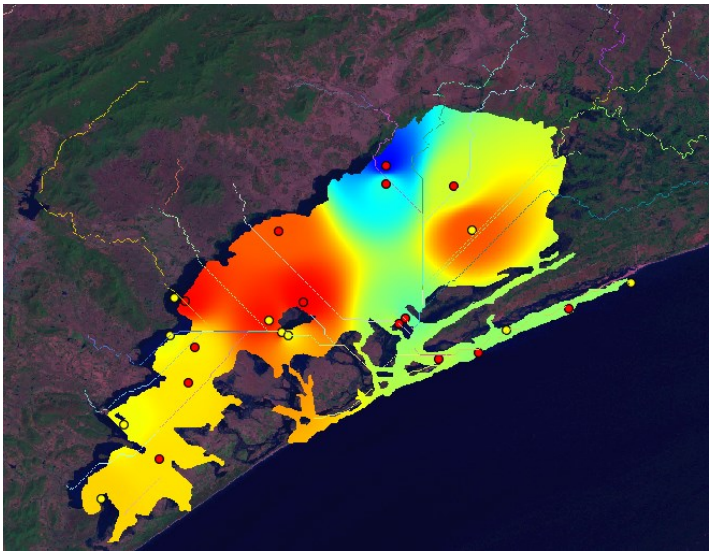
FishFarm



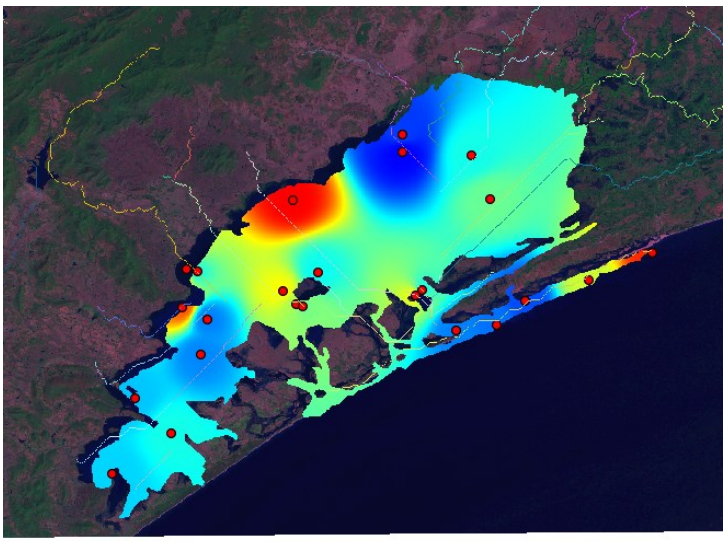
pH - Rainy season



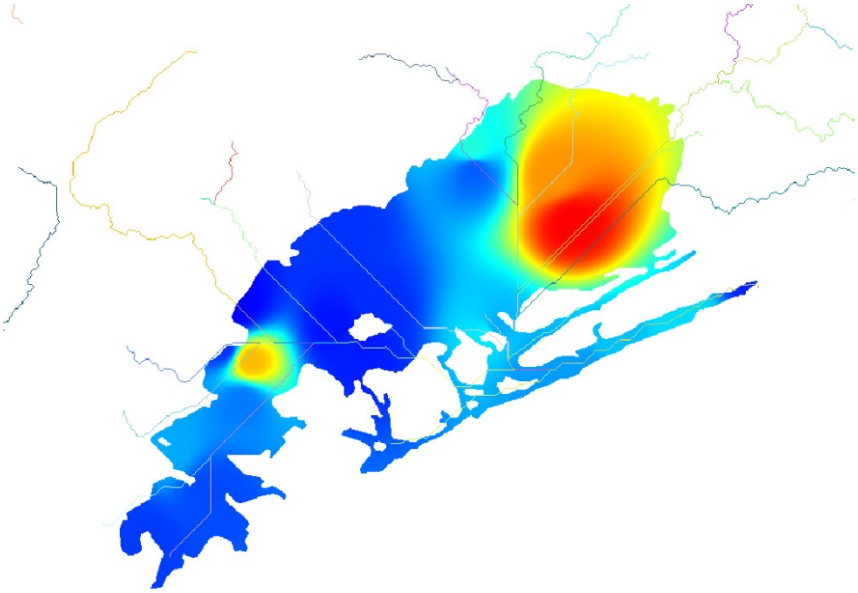
pH summer season



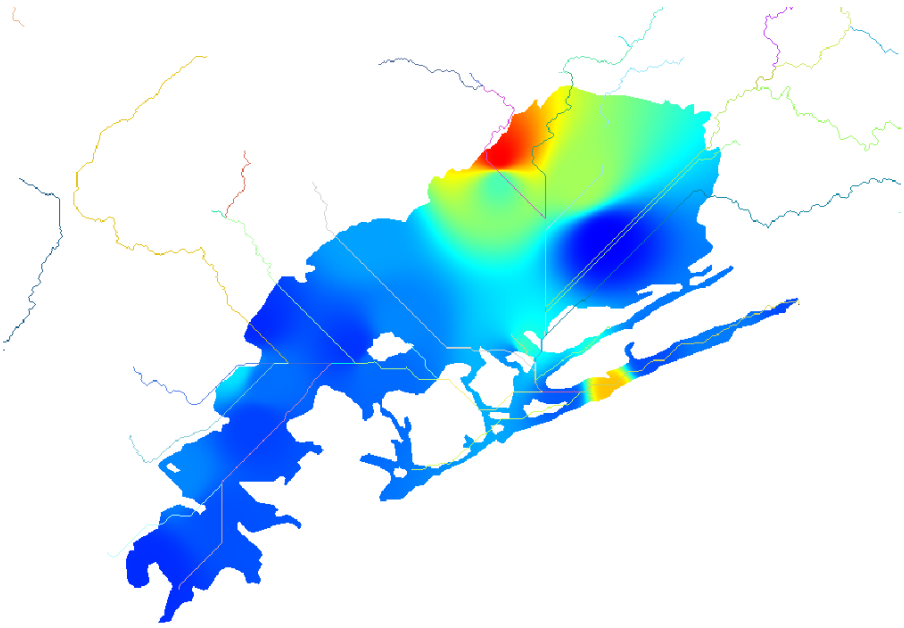
pH winter season



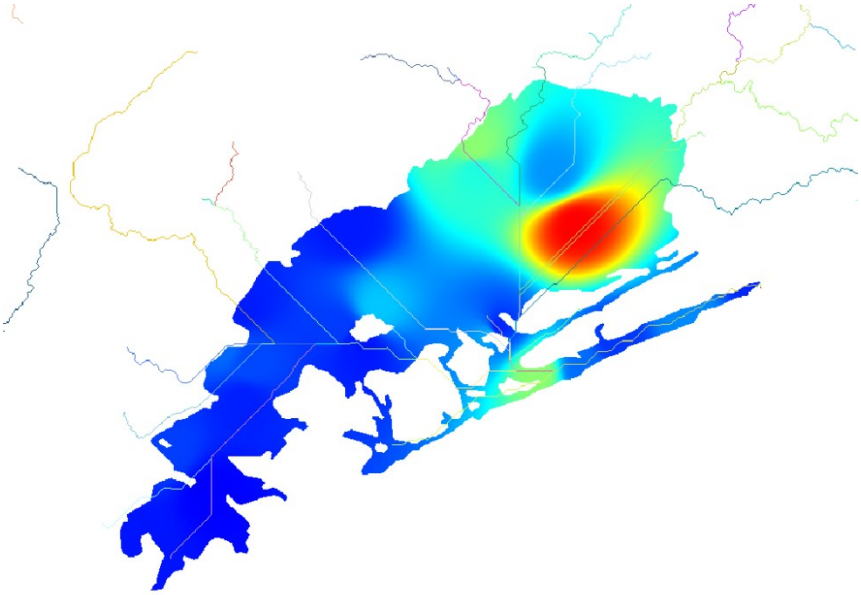
Turibidity summer season



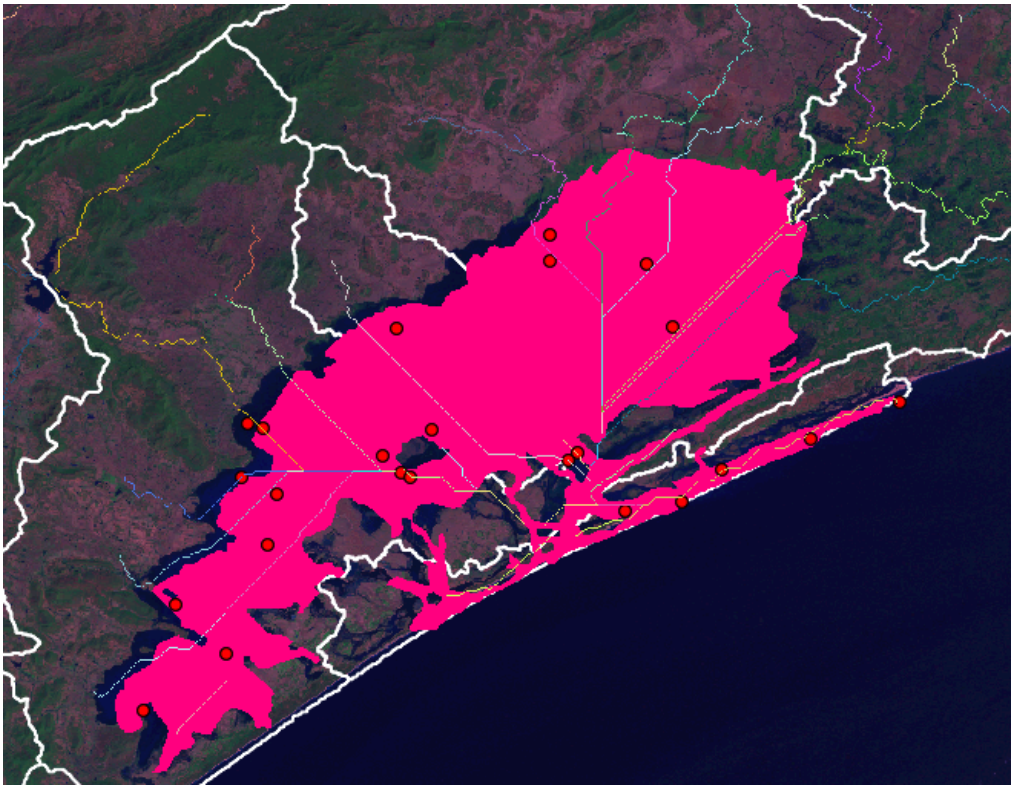
Turbidity Winter season



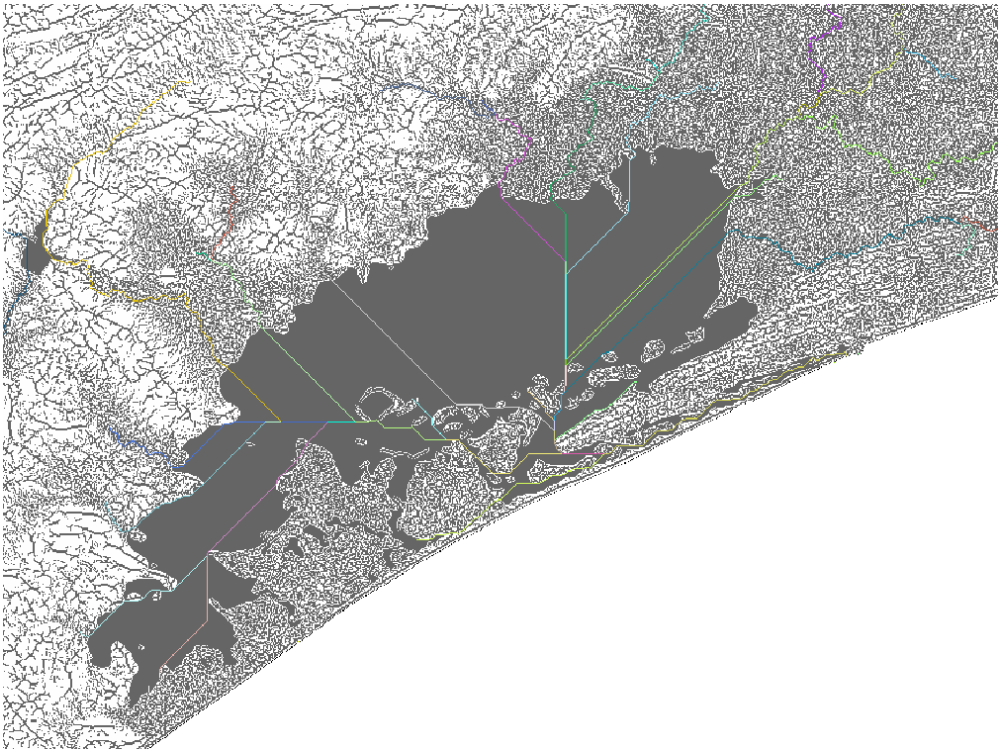
Turbidity Rainy season



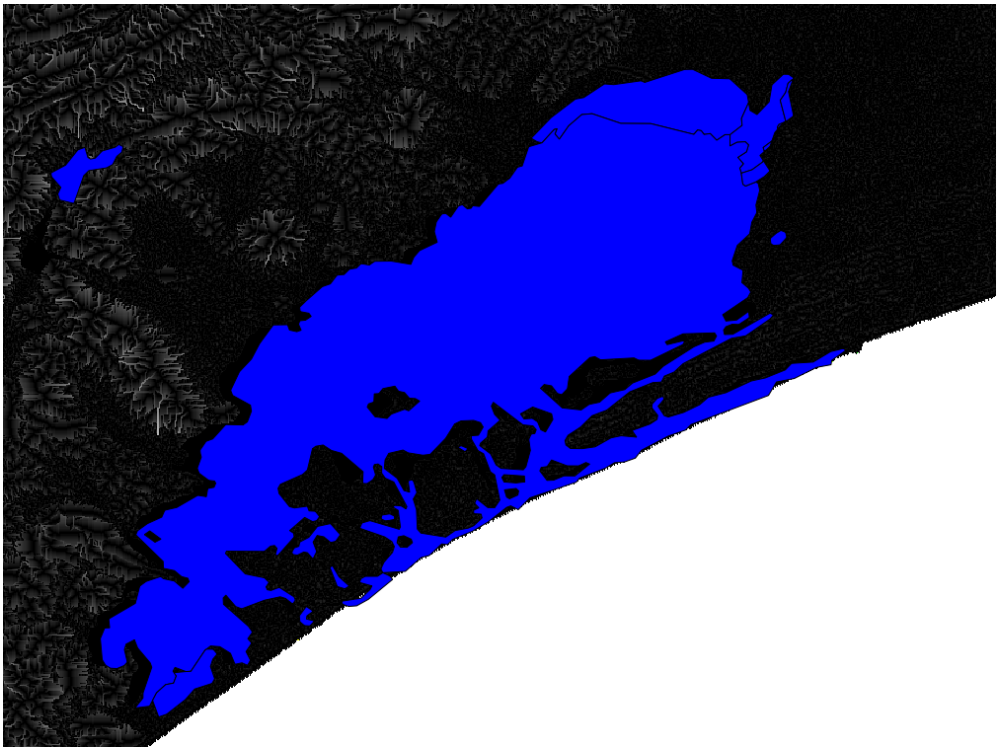
Large Watershed



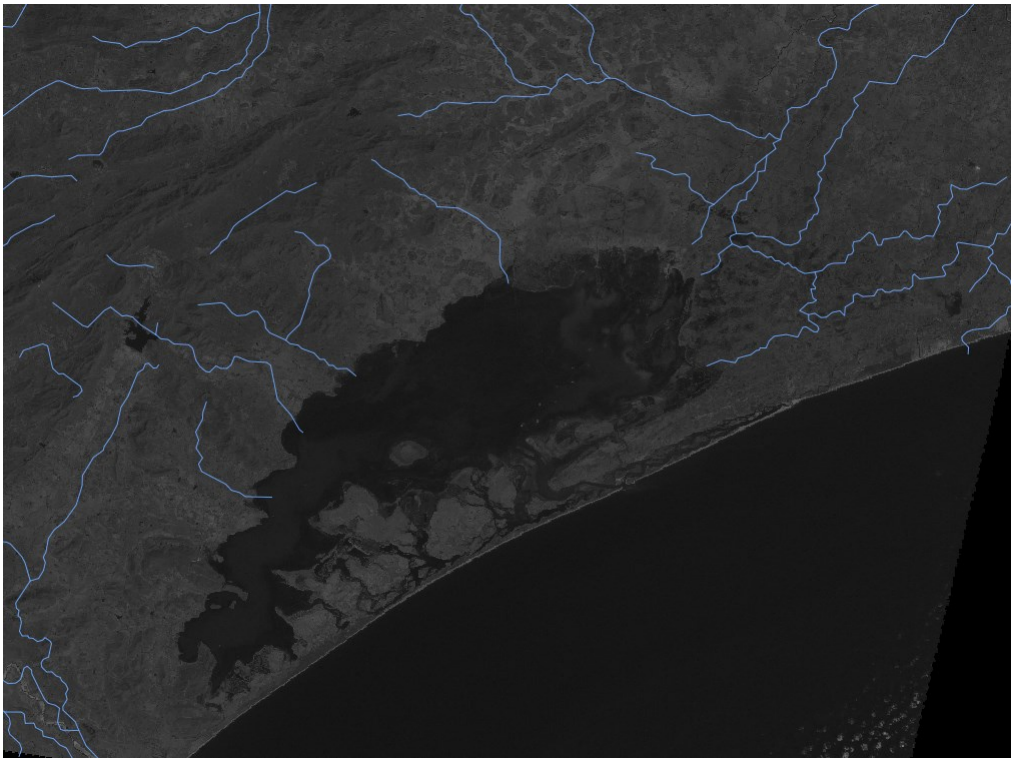
Slope steepness



Slope Lenght



Rivers



.